

PROFESSOR McDougall's Contributions to The Science of Psychology⁽¹⁾

By HERBERT S. LANGFELD

Princeton University

CONTRARY to common opinion great men have usually been precocious in their youth; and McDougall was no exception. Before his first graduation in science from the University of Manchester at the age of 17, he was already acquainted with the writings of the leading scientists of the time. Not satisfied with having received first class honours at this provincial university, he entered St John's College, Cambridge, on a scholarship and continued to distinguish himself in his studies. During these eight years of intensive college work he laid the foundation for that vast learning which was evident in all his writings. At Cambridge he took a medical degree, not with the intention of practising but because he considered a medical training necessary for one desiring to study human behaviour scientifically. In this early conviction we see the germ of that endeavour to raise physiological psychology to the position of a dignified and productive science in England. That he did not remain entirely in physiology is probably due in part to the influence of William James and in part to something intrinsic in his own personality which caused him always to strive for the most inclusive view of human nature. That he should have been influenced more by James than by any other psychologist of the time is not surprising, since they had many attributes in common. They were both men with broad, cultural backgrounds. They had vivid and constructive imaginations, which enabled them to grasp important relations which often escaped the ordinary scholar. They had a contempt for those who split hairs over petty facts, with the result that their energies were always directed toward the most fundamentally important problems. They had a common dislike for the pedantic and purely academic, and were bored by the dry systematic German psychology. Finally, they both had a keen aesthetic sense, which was evident in the smoothness, fluency and lucidity of their literary style.

McDougall's first publication at the age of 26 was on a purely physiological problem—the structure of the cross-striated muscles and the nature of their contraction⁽²⁾. A year later, however, he started to

publish in the field of psychology, his first paper being upon psychological method (3).

1901 must have been an exciting year for the young scientist. He had the privilege of reviewing two works which had a profound effect on experimental psychology—Müller and Pilzecker's experiments on memory, and Titchener's *Experimental Psychology*. More important for him, however, was the publication of the report of his excellent work on vision (4), which marked the beginning of an interest in that field which lasted the rest of his life. He had first become acquainted with Hering's work during his undergraduate days through Rivers, who had studied with Hering in Germany. He soon found himself in disagreement with Hering's theory and decided to put the hypothesis to further experimental test. These were still early days in experimental psychology and the instruments were necessarily crude. All that McDougall had at his service were coloured glass and gelatine filters, and when he could not use sunlight, he had to resort to the flame of an oil lamp. Yet in this paper he reported the results of fifty separate experiments. As one reads the description of this extensive research, one is impressed with the keenness of McDougall's observation, with the ingenuity with which he varied the conditions in order to check the various assumptions of the existing colour theories, with his wide knowledge of the literature and his appreciation of the relation of his results to other aspects of human behaviour. He suspected the weakness of the Hering theory from the fact, as pointed out by Mrs Ladd-Franklin, that "approximately pure or simple" red and green are not complementary. Therefore, through experimentation, he proved to his own satisfaction the nature of colour mixture, and swung around to the Young-Helmholtz theory, or rather—through patriotic reasons, perhaps—to the Young version. He stated: "Young's theory, . . . if we incorporate with it the hypothesis of the separate W-exciting apparatus having its retinal seat in the rods, as urged by v. Kries, Rollet's theory of contrast and the theories of induction and after-images set forth above, explains fully all the phenomena described above; and, with the exception of certain observations of a complex character, very difficult to interpret, such as Müller's experiments on galvanic stimulation of the eyes and some of Prof. Sherrington's flicker experiments, I cannot discover any important fact connected with light- and colour-vision with which it is incompatible or that it fails to illumine." (5)

He also studied the mutual inhibition of visual images, and undoubtedly inspired in part by Münsterberg's action theory to which he

refers, he examined the reinforcing influences of the activity of the ocular muscles. He saw the importance of the part played by the cortex in visual phenomena, and his eager mind went beyond colour theories to the more fundamental concept of inhibition.

McDougall always thought and wrote rapidly (though curiously enough he spoke slowly and hesitatingly, a characteristic due no doubt to his self-consciousness in social situations), and he was therefore able to publish a surprising amount of material. In the same year appeared his paper "On the seat of the psycho-physical processes" (6), a theoretical discussion based partly on his observation in the visual experiments. In it he argued that the synapses are the most readily fatigued parts of the cortical conduction paths, and that the amount of fatigue is related to the amount of neurin discharging across the synapse, and thus to the degree of potential in the system. The drainage theory, we see, was already beginning to take form in his mind.

In the year 1902 he wrote the first part of his famous paper, "The physiological factors of the attention process" (7), in which his drainage theory was described. In it he exhibits two trends which have always characterized his speculation, the one toward explanations in dynamic terms, and the other toward physiological principles. The drainage theory maintained that the simultaneous excitation of two nerve paths increases the intimacy of the connexion between the paths and facilitates the flow of neurin into the path most strongly activated.

Voluntary attention gave him trouble. He stated: "...while all attention involves concentration or convergence of free nervous energy from all or many parts of the brain into some one neural system, voluntary effort results in a further degree of this concentration of energy" (8). The increase of attention through voluntary effort, he felt, could not be explained through physical causation alone. He was not satisfied either with Münsterberg's or Ribot's theory of muscular response. Such responses in voluntary attention were not to be denied, but they were merely secondary manifestations of attention. Now McDougall was always an unwavering believer in the interaction of the mental and physical, a position which he most ably defended in the classic *Mind and Body*, and he felt that the psychologist must seek "in the insufficiency of physical cause evidence of psychical efficiency" (9). He therefore intimated that in voluntary attention we may have evidence of psychical guidance of physical processes. His drainage theory had accounted for the simpler manifestations, but even in this early period of his career such an explanation was too mechanistic to account for many of the

110 *McDougall's contributions to the science of psychology*

higher processes of mental activity. The purposive aspect of life was taking hold of him. He pushed the mechanistic explanation as far as he thought it consistent with experience, and would then seek the aid of teleology. Yet he did not minimize the importance of the strictly causal relation of the physiological processes.

McDougall continued to work on visual phenomena and in 1903 published further data on the inhibitory processes (10). Again his equipment was simple—a spinning top, some visual patterns—but his conclusions were far-reaching. The observed brightness and colour contrasts, he felt, must be produced by inhibition, but existing theories such as Hering's, Verworn's, Wundt's and Gaskell's were inadequate. It seemed evident to him that the more intensely aroused processes tended to weaken the less intensely aroused through the process of drainage. But his scientific imagination did not stop here. In his own words, "...any sensory presentation or idea in becoming the object of attention, the central object of mental activity, causes any other object or idea to cease to be attended to. It is, therefore, clear that there obtains a relation of reciprocal inhibition between all those numerous organised systems of neurones or nerve-paths, the physiological bases of the mental systems, which have been built up through successive apperceptions (partly of the individual, partly of the race) to bring about appropriate reactions to every kind of object presented to the senses, or, as we say when the reaction is complex, appropriate conduct in relation to the object" (11).

Shortly after this paper appeared, 1904, the time seemed ripe for a psychological journal in Great Britain. The publication was named *The British Journal of Psychology*, and the editors were Ward and Rivers, the one a theorist, the other primarily an anthropologist. Seven of the nineteen papers of the first volume were experimental; five of them bore the name of William McDougall. We find from these articles that he was still primarily interested in visual phenomena and inhibition. The phenomenon of Bidwell's ghost he explained by an ingenious theory of intermittent nervous discharge, faintly suggestive of the refractory phase of nervous conduction. The fluttering heart illusion, that interesting experimental demonstration of the first decade of the century, gave him further ammunition with which to attack Hering and to support v. Kries's duplicity theory. From the results of his experiments on the latency of response of different wave-lengths, he concluded that there was little difference in their 'action times' if the colours were made to appear of equal intensity.

He accepted Helmholtz's explanation of the predominance of contours in the struggle between two visual fields as due to attention, but as we would expect he desired to describe attention in terms of the competition of nerve paths for the supply of free nerve energy. It is interesting to note his suggestion that his drainage theory furnishes an explanation of Sherrington's reciprocal innervation and inhibition.

In the same number of the *British Journal* he also described a method of studying concurrent mental operations and the resulting fatigue. His method was for the subject to mark dots which appeared on a revolving drum, the hand movements being recorded graphically. An arrangement similar to this became a familiar set-up in experimental laboratories and in work in the applied fields.

His *Physiological Psychology* (12) appeared in 1905. This little volume had only nine short chapters, four being on physiological processes and the others on perception and instinct, emotions, retention (including drainage), learning, and voluntary and ideomotor action; but the book was a godsend to the teachers of experimental psychology of those years that directly preceded the appearance of Charles Myers' *Textbook of Experimental Psychology*.

In 1908 he published with Hocart the results of his observations on auditory localization (13). He accepted intensity differences as the cue for localization, but as he believed these differences were often subliminal he favoured a local-sign theory. He did not fail to observe, however, that noises and complex tones are localized with surprising accuracy, a fact which he confessed was hard to explain by means of his theory, but which he probably would not have found so baffling if he had had the present instruments for regulating the phase, complexity and intensity of the tonal stimulus.

McDougall's active mind was always finding interesting situations for experimentation. In this same year he reported the results of his investigation of the colour sense of two of his children (14). He used a modified arrangement of Baldwin's method of having the infant reach for one of a pair of coloured objects, and found that the hues red, green and blue were appreciated at the age of six months, but that there seemed to be no real colour preference. The following year, 1909, he used his theory of fatigue, which he believed due to increased resistance in the synapses, to explain both neurasthenia induced by shock, and hysteria (15). In 1911 he applied his principle of drainage to the phenomenon of retinal rivalry (16), and in 1914, together with Flugel, he published "Some observations on psychological contrast" (17). He experimented upon the judge-

ments of lifted weights, judgements of length of lines, of width of angles, of duration, and of the rate of seen motion. The evidence he obtained inclined him to accept Müller's theory of motor-set with reservations. He never really felt satisfied with an explanation entirely in behavioural terms, and so we are not surprised to find that he added perceptual set to motor-set.

Then came the war and his scientific work was seriously interrupted. In 1920 he appeared in a new field, that of learning and retention, when he reported a series of experiments made with the co-operation of Miss May Smith⁽¹⁸⁾. The results convinced him that practice in memorizing produces both improvement in memory and in ability to commit to memory. In short, memory, as is widely recognized to-day, seemed to him to be a dynamic process and not the effect of mere repetition.

1926 saw him return to a defence of the drainage theory against an attack by Raymond Dodge⁽¹⁹⁾. He admitted in this paper that he had been so long out of touch with physiology that he felt quite incompetent to discuss the topic. However, after reading Anrep's translation of Pavlov's work, he wrote in 1929, "...all of the many phenomena of inhibition so brilliantly studied and reported in this great work seem to lend strong support to the hypothesis of inhibition by drainage which I formulated 25 years ago..."⁽²⁰⁾.

In his student days at Cambridge University he rebelled against the prevailing and all-dominant theory of Weissmann which repelled him by its dogmatism. In 1920 while at Harvard he started that long series of experiments on successive generations of white rats in an endeavour once and for all to prove or disprove the Lamarckian hypothesis. In his last published paper on the experiment he wrote: "The future of this experiment is uncertain. I am an old man and a cripple, and cannot foresee the duration of the opportunities for prosecution of it which I still enjoy at Duke University"⁽²¹⁾. Fortunately he will not have laboured in vain, for his son, Kenneth, is continuing the work.

Time does not permit a review of his extensive work in the abnormal field which resulted in one of the best text-books on the subject, or of his great interest in psychic research and kindred phenomena. I trust, however, that this brief account of his enormous productivity will convey some idea of his genius. Certainly his work speaks more clearly than anything I could say.

McDougall was a very firmly integrated personality. The pattern of his scientific thought was woven very early in his career, and he never broke one of its threads. If he accepted a theory, or formulated a new

one, he defended it patiently and fairly against all attacks. The best example of this characteristic is his defence of the drainage theory which has been questioned by most of the eminent physiologists, and his rebuttals to the many criticisms of his Lamarckian experiments. He always believed in the interaction of the mental and the physical, and vigorously attacked Watsonian behaviourism when it was at the height of its popularity. He believed in purpose as a directive force and criticized most of the prevailing theories of learning, even that of Tolman. He agreed with the Gestalt school in its emphasis of the total situation, but the research which he did with his son, Kenneth, on the monkey, racoon, rat and wasp led him to assert that insight is not a sufficient explanation of animal behaviour. Until the Gestalt psychologists accept a conative guide by foresight, he said, their theory will remain mechanistic in the broader sense of the term.

McDougall distrusted any doctrine that was popularly accepted. He was therefore constantly pulling against the stream, but he fought with the faith of a crusader and with an indomitable courage and unswerving purpose that commanded the respect of all who knew him.

REFERENCES

- (1) Based on a paper read at the Memorial Service for Prof. William McDougall, Duke University, Durham, North Carolina, 7 April 1939.
- (2) "On the structure of cross-striated muscle, and a suggestion as to the nature of its contraction." *J. Anat. Physiol.* (1897), xxxi, 410-41.
- (3) "A contribution towards an improvement in psychological method." *Mind* (1898), vii, 15-33, 159-78, 364-87.
- (4) "New observations of Thomas Young's theory of light- and colour-vision. I-III." *Mind*, N.S. (1901), x, 52-97, 210-45, 347-82.
- (5) *Idem*, p. 382.
- (6) *Brain* (1901), xxiv, 577-630.
- (7) *Mind*, N.S. (1902), xi, 316-51; (1903), xii, 289-302, 473-88; (1906), xv, 329-59.
- (8) *Idem*, p. 359.
- (9) *Idem*, p. 359.
- (10) "The nature of inhibitory processes within the nervous system." *Brain* (1903), xxvi, 153-91.
- (11) *Idem*, p. 168.
- (12) *Physiological Psychology*. 1905. London: Dent (New York: Macmillan). Pp. xiii + 172.
- (13) "Some data for a theory of the auditory perception of direction." *Brit. J. Psychol.* (1908), ii, 386-405.
- (14) "An investigation of the colour sense of two infants." *Brit. J. Psychol.* (1908), ii, 338-52.
- (15) "The conditions of fatigue in the nervous system." *Brain* (1909), xxxii, 256-68.

114 *McDougall's contributions to the science of psychology*

- (16) "On the relations between corresponding points of the two retinae." *Brain* (1911), xxxiii, 371-88.
- (17) *Brit. J. Psychol.* (1914), vii, 349-85.
- (18) "Some experiments in learning and retention." *Brit. J. Psychol.* (1920), x, 199-209.
- (19) "The hypothesis of inhibition by drainage." *Psychol. Rev.* (1926), xxxiii, 370-4.
- (20) "The bearing of Professor Pavlov's work on the problem of inhibition." *J. Gen. Psychol.* (1929), ii, 232.
- (21) "Fourth report on a Lamarckian experiment." *Brit. J. Psychol.* (1938), xxviii, p. 393.

THE 'TRIVIAL' AND 'POPULAR' IN PSYCHOLOGY¹

By T. H. PEAR

From the Department of Psychology, University of Manchester

- I. *Academic compared with non-academic 'triviality' and 'popularity'* (pp. 114-117).
 - II. *'Trivial' items in the dream and the psychoneurosis* (pp. 117-118).
 - III. *'Trivial' must be defined in some frame of reference* (pp. 118-119).
 - IV. *Comments upon some recent psychological publications* (pp. 119-121).
 - V. *Social psychologists' neglect of speaking* (pp. 121-122).
 - VI. *Do some branches of psychology afford escape for the introvert?* (pp. 122-123).
 - VII. *'Popular' psychological subjects* (pp. 123-124).
 - VIII. *Fifty years of change in American psychology; Prof. G. W. Allport's summary* (pp. 124-127).
 - IX. *Dangers of pseudo-austerity in the psychologist's mind* (p. 127).
- References* (pp. 127-128).

I. ACADEMIC COMPARED WITH NON-ACADEMIC 'TRIVIALITY' AND 'POPULARITY'

"'TRIVIAL'", says Webster's *Dictionary*, "is derived from Latin *trivialis*; that belongs to the cross-roads: hence, common, from *trivium*, a place where three roads meet." The first meaning, 'ordinary, common-place, trite', is said to be obsolete or archaic. The second meaning given is 'trifling, petty, paltry'.

'Popular' means (1) pertaining to the common people; (2) plebeian, vulgar; (3) suitable to the public in general as (a) easy to understand, plain, as *popular science*, (b) adapted to the means of the common people, hence, cheap, as *popular prices*; (4) believed or approved by the people, as a *popular leader*. It is perhaps permissible to doubt whether meaning (2) is obsolete among psychologists in Great Britain.

In this article it will be contended that (a) in this country and in America during the last twenty years, many psychologists have avoided the 'trivial' and 'popular' so consistently that their studies are rapidly veering away from many valuable contacts with real life; that (b) in eschewing so-called 'triviality' they have created, developed and upheld

¹ Substance of a paper read to the British Psychological Society's Birmingham meeting on 5 April 1940.

categories of academic triviality, which threaten to become vested interests; that (c) avoidance of 'popularity', either as persons, speakers or writers, with the common people, is often linked with a fear of unpopularity with the ruling powers of science, especially of the physical sciences, and that (d) these tendencies threaten to dehumanize, emasculate and stultify psychology in general.

Some may object here that psychologists have no more right than anyone else to judge the triviality of a subject, for 'trivial' describes something which is to be compared with other things, against a context, background, or frame of reference. This is partly true. An insect bite is trivial if it does not happen to poison its victim—and to make this judgement needs no psychological insight. Historical instances are numerous where 'trivialities' have proved serious. Though historians love 'debunking' legends, we may perhaps believe that Captain Jenkins's ear and the alleged presence of pig lard in the Indian cartridges played some part in a war and a mutiny respectively. Yet some critics of psychology confidently set themselves up to judge what is 'trivial' in this study. They may be right, but I suspect that they are more firmly embedded in their own culture patterns than any psychologist ought to be.

Investigators employed by the British Institute of Public Opinion show how difficult it is to believe that at present our rulers, even our own political representatives, keep themselves well informed concerning many social matters which the majority of citizens consider to be of major importance. This fact has a practical as well as a theoretical aspect. Prof. J. D. Bernal⁽¹⁾ and Prof. Lancelot Hogben⁽²⁾ maintain that the common acceptance of the point of view of the well-fed, comfortable, academic man leads to a state of affairs in which research into human relations, particularly in sociology and economics, is publicly tolerated only in so far as it has futility value, for if many of the recommendations arising out of these studies were put into practice, the economic repercussions would be considerable, and serious jars would be given to the delicate stratification of our social system; a stratification which, by the way, in its upper and middle layers, is mirrored in some English universities. We glimpse here the possibility that a subject which many academic people call trivial may, for that very reason, sometimes be regarded as having established a *prima facie* case for serious consideration by the practical minded.

An example of 'unacademic' university research was made famous by Dr Abraham Flexner in his *Universities, American, English, German*⁽³⁾. A thesis for the M.A. degree in the University of Chicago was entitled

"A time and motion comparison of four methods of dish washing", and Dr Flexner made merry with this fact, causing titters in university common-rooms for years afterwards. Now if the actual investigation of dish washing had been observed in detail—I neither affirm nor deny that it was so observed—by Dr Flexner, it might have proved to be the reverse of trivial, though still, alas, popular. Dishes and the processes of cleaning them are common enough, but presumably the research aimed at understanding these processes and, consequently, improving them and eliminating from them unnecessary or fatiguing movements, or manipulations likely to break the pottery. In restaurants and canteens large-scale dish washing has been studied by the National Institute of Industrial Psychology. Serious breakages, both of china and of friendly relations, were substantially reduced by research into their causes. Moreover, dish washing lends itself to 'motion study', which, since it occasionally requires complicated and expensive apparatus, may appear scientific even to a natural scientist. In any case, the principles underlying motion study are significant for the physiologist or psychologist; a matter discussed by Frederick W. Taylor and Lillian M. Gilbreth, C. S. Myers and E. Farmer. Motion study may be admitted to have economic importance, since there have been serious and prolonged strikes about its use. Yet Dr Flexner submitted the 'list of Chicago theses' of the Department of Home Economics and Household Administration, including, presumably, the dissertation we are considering, to "two eminent authorities in the field of biochemistry and nutrition", for their opinion.

It may be objected that Dr Flexner's contempt was aroused by the fact that such a thesis was submitted for the Master of *Arts* degree, or that he was criticizing adversely the method, not the subject, of investigation. Be that as it may, in my experience, when English university teachers have deplored, sometimes rather pharisaically, the 'triviality' of themes for doctorate theses, in the U.S.A., dish washing has been the *leitmotiv*. So I think Dr Flexner's example may have been unfortunately chosen, even for his purpose.

II. 'TRIVIAL' ITEMS IN THE DREAM AND THE PSYCHONEUROSIS

But we have not merely to discuss whether Smith's idea of a 'trivial' subject is better than Jones's, and, inevitably, decide that there is no quarrelling about tastes. The problem is subtler than this. It is now forty years since Freud first maintained that the dream does not bother itself with trifles. He taught, moreover, that in the manifest content of a

dream, an apparently trivial item may symbolize something very significant in the latent content, and that in waking life an apparently 'trivial' liking or phobia may do precisely the same thing. This is true of some 'autistic gestures', and of many instances of mildly pathological, even of healthy behaviour.

III. 'TRIVIAL' MUST BE DEFINED IN SOME FRAME OF REFERENCE

To use a flashlight momentarily in a darkened room with the blinds up may be a trivial act in the 'black-out', but in peace time this action would arouse justifiable suspicion in any policeman. It is a trivial act for an adult to enter his house after midnight, but not if he be an undergraduate in Oxford or Cambridge. The ceremonial of baptism; changing the guard outside a royal palace, the thousand and one details of the last Coronation are trivial regarded from many points of view, but not in the eyes of those who insist upon them. I believe that the almost non-existent interest of most psychologists in problems of clothing and speaking, and of fashions in them, is due to the fact that in universities both these subjects are regarded as trivial, except in the examples of wearing sports' millinery or speaking a foreign language. Yet in the everyday world, the man who has failed to secure an important appointment because of his unsuitable clothes or speech, and the woman whose small business is ruined by a sudden change in fashion, will regard these matters as tragically significant. The British War Office has found that such a trivial gew-gaw as the kilt seems important to many people, who wish it to be retained even when its wearer may be exposed to irritant gas or lice.

Psychologists, of all people, ought to be interested in the ways by which public opinion is formed. The apparently 'trivial' is often deliberately exploited by the publicist. Incessant reiteration in the newspapers and on the radio of 'trivial' facts about little-known people, is a well-known method of building up a 'public attitude', which will ultimately cause these persons to be regarded as important. More obviously useful is the technique of 'plugging' trivial news about important personages. Who believes, for example, that the numerous 'trivial' details about the lives of the Royal children had no effect on the average person's opinion of the present King and Queen?

'Trivial' must always be defined with regard to some value. This may be temporary and unusual, though none the less real and cogent. That on the morning I wrote this I was feeling cold is a trivial fact if you compare

my experience with that of millions to whom the snow was causing agony. Yet if, being too cold to think about this article, I fail in my duty to my university or my family, the cold is trivial neither to me, nor to them.

Now for the obvious. Again and again, literary critics, nourished perhaps upon the wilder imaginings of some camp-followers of psycho-analysis, condemn a psychological book as containing 'obvious' facts about human experience or behaviour. Yet a textbook of zoology which omitted to mention, let alone describe, the obvious horse, dog, or frog in order to seem fashionably knowledgeable about pandas and okapis, would be condemned by scientists. Moreover, in psychology many 'obvious' statements are true only of a particular type of mind; for example, confident assertions, by non-visualizers, about the limitations of visualizers, or, to an introvert, the extraverted behaviourist's generalizations based upon his own often unnecessary meddlings with the outside world. It is safe to assert that to-day psychology requires many more careful records of unexciting events and facts. In this connexion, attacks upon what 'mass-observation' workers call 'academic psychology' seem justified, for many ways of behaviour, obviously important in a university, may have no parallel in Worktown, and seem trivial to its inhabitants.

IV. COMMENTS UPON SOME RECENT PSYCHOLOGICAL PUBLICATIONS

At the moment, I am puzzled by certain references to so-called 'trivial' matters in psychology. A recent effort by me⁽⁴⁾ to describe some psychological problems of everyday life appeared to a reviewer⁽⁵⁾ "at first sight popular and at times even trivial, but a second reading reveals considerable thoughtfulness in its preparation". This chapter, rewritten several times, and submitted for criticism to about a dozen social psychologists, treated of the following subjects: friendly relations between nations, social stratification, clothing, cosmetics, speaking, propaganda, the cinema, the radio, 'listener research', television, personality, unemployment, the attitudes of different races and nations towards each other, religious belief, conversation and war.

At the risk of being tedious, I will examine this reviewer's attitude, for I believe that the fashion among psychologists of not answering criticism, though it may preserve a rather queer form of personal dignity, is bad for psychology. To whom are the above-named subjects trivial? Possibly to a Trappist monk, a schizophrenic, and an old-fashioned type

of university don, but to whom else? Is it the mention of clothing and cosmetics (a form of clothing) which shocked the critic? Yet can a psychologist who lives in the biggest textile area of the world, forget the connexion between clothing, economics, and politics? Surely there is social importance in the fact that nowadays to put on cosmetics will lose a woman one job, while to appear with an achromatic face will lose her another? And is there any connexion between personality, vocational fitness, economic success, individual happiness and social utility? One would have thought so.

If I mention another less severe review (6) of the same chapter, I fear that in England I may be regarded as having the bad taste to argue with the umpire. But I do this because in neither (unsigned) review has the judge indicated his own opinions; and those, at the present moment, might have helped me greatly. The reviewer says:

Prof. Pear puts greatest emphasis on 'the social psychology of everyday life'; problems connected with unemployment, broadcasting, the cinema, propaganda and conversation loom largest amongst his suggestions. He has selected problems which he regards as 'especially important for psychological theory or practice', but unfortunately, he has not indicated at all fully where their psychological importance lies. As a result, his suggestions tend to seem arbitrary and at times verging on the trivial. This is no doubt due to the brevity of the article, and one wishes that Prof. Pear had spent longer in developing the wider implications of his problems and placing them in *proper*¹ perspective.

But as I see it, if I assert that psychologists are neglecting important problems of everyday life, and specify them, I am condemned as arbitrary. I will therefore become more assertive, and say that in the *Handbook of Social Psychology*, edited by Carl Murchison (7), a book of 1158 pages, the social problem of clothing might have been given rather more than two; about one-sixth of 1 %. The frivolous topic of cosmetics is actually mentioned here, though in a discreetly anthropological manner which will not harm the author's reputation in any university. The passage (which refers to clothing as a social custom) is this: "If a bit of mud is smeared on purposely, that would fall under this type of behaviour." Dr Clark Wissler concludes these two pages by adding (p. 550): "Space forbids further analysis of clothing and dress in its material, social and aesthetic aspect." And though this huge book is dated 1935, Prof. J. C. Flugel's *The Psychology of Clothes* (8) is not mentioned. Yet this *Handbook of Social Psychology* devotes 310 pages (27 % of the whole) to the social phenomena of bacteria, plants, insects,

¹ Italics mine.

buds, mammalian herds and packs, infra-human primates, and the physical environment. The recent *Introduction to Psychology* (not, be it noted, merely to Experimental Psychology) by Boring, Langfeld & Weld(9), contains no reference to the undoubtedly popular (and therefore trivial?) subjects of War and Peace—and it was published in 1939. I feel that the verdict of the general public might be that my selection of subjects in that chapter was less arbitrary than the choices quoted above.

V. SOCIAL PSYCHOLOGISTS' NEGLECT OF SPEAKING

I have already indicated that the study of clothing is neglected by most psychologists. A glance at the index of books of psychology will suggest that there is a similar taboo against the discussion of speaking. Probably the word 'taboo' is applicable here in its severely technical as well as its most popular sense, for while many thinkers are deeply interested in clothing and speaking as signs of personality, they are often reluctant to discuss them with reference to a particular person. I will challenge a reviewer once more, and once more only. In a notice of the present writer's *The Psychology of Effective Speaking*(10), the anonymous judge records that the book would be improved if it were "tricked out with fewer irrelevancies"¹ (no examples given), but adds: "A plea that the speaking of English should be taught in the schools deserves attention, however." Is this considered an almost irrelevant subject? The 'however' suggests it. But are the chapter headings in this book about subjects irrelevant to public speaking? They include: an examination of lectures and lecturers, particularly in universities, a discussion of spoken and written style, prejudices for and against dialect, a consideration of speech reform in view of the rapidly increasing importance of radio—admittedly a 'popular' subject, but hardly trivial. Were these irrelevances? Or was the chief irrelevance a chapter on the uses of humour in speaking? Possibly, for while it seems an axiom that psychologists may write of laughter, they may not, without losing caste, write about humour. Even Freud did not venture past a consideration of the joke.

At this point, some may try to dismiss the subject, saying that it is useless to dispute about tastes. I contend, however, that to-day it is at least useful to discuss them. Many psychologists to-day are, I submit, turning away from important psychological problems, and since social forces are largely responsible for this behaviour, discussion of them is significant and relevant, if unpopular. Indeed, it belongs to that neglected

¹ 'Irrelevant' joins 'trivial' and 'popular' as an epithet which expresses one's own views about psychology's 'frame of reference'.

study, the psychology of psychologists. I beg to be allowed, in the next page, to utilize material from a recent article (11).

As Prof. J. D. Bernal points out in *The Social Functions of Science* (1), scientists have lost their amateur status. Most psychologists nowadays are completely professionalized; a fact which inevitably influences the directions along which their science is developing. Study of the sociological setting of psychologists' activities, therefore, promises valuable results, for the development of all the sciences has greatly depended upon the idiosyncrasies of prominent workers, and upon the economic and social encouragement they succeeded in winning.

VI. DO SOME BRANCHES OF PSYCHOLOGY AFFORD ESCAPE FOR THE INTROVERT?

In his urbane summary of *A Hundred Years of Psychology* (12), Prof. J. C. Flugel, illustrating Dr Alfred Adler's belief that one's original inferiority in some respect may be converted into a superiority by constant exercise and effort, remarks:

as when Demosthenes, the stammerer, became one of the greatest of orators, when Sandow, the weakling boy, became a recognized strong man of his day, or again, when a person gifted with exceptionally little insight into the thoughts and motives of his fellows compensates by becoming a psychologist.

I take this remark quite seriously. Not all psychologists like human beings more than anything else, or handle them unusually well. This raises some important questions. Are 'shy' psychologists common? Is their number increasing? If so, how will this influence the theoretical and practical development of the study?

In remarking that some psychologists like people, it is recognized that a few may love their fellow-men too much to judge them impersonally. Yet some forms of psychological judgement—perhaps all—are impracticable without sympathy, since the psychologist's task is to observe minutely, and describe accurately as well as to explain satisfactorily. To-day, for excellent reasons, psychologists are often immigrants from the older sciences. They may bring with them some attitudes common in these studies, including a distaste for intimate personal acquaintance. Connected with this fact, perhaps, is the claim that the study of mathematics purges the mind of sentimentality—even if in so doing it implants an over-ruling sentiment for mathematics.

I contend that a laboratory may be a refuge for thinkers whose affections for human beings is subnormal. Unsociable youths may leap

at a chance to study man from the eyrie—or the ivory tower—of a laboratory instead of face to face in the real world, where they might be expected to show enthusiasm or sympathy, even to risk making fools of themselves, or appearing cheap in the eyes of austere scientists. This fear, therefore, is a sufficient reason for avoiding the ‘popular’. I would go further, and ask, are modern psychologists afraid even to be interesting? Is there not an almost puritanical tendency causing many psychologists to avoid studying subjects of undoubted popular interest? This austere aloofness is not the prerogative of psychologists, yet together with anthropologists and sociologists, they have recently been criticized for neglecting their proper studies. The Institutes of Public Opinion in America and England, by their very choice of subjects to investigate, and the lively, even noisy, mass-observation movement in England, suggest that professional students of the human mind are fiddling while Rome burns.

In fairness to both sides, it might be pointed out that the day before these lines were written, *The Study of Society* appeared over the names of seventeen of the reproached ones. Yet many of these would readily admit that a large number of important and popularly interesting subjects are barely mentioned in this book.

VII. ‘POPULAR’ PSYCHOLOGICAL SUBJECTS

The meanings attached privately by academic people to the word ‘popular’, if examined closely, might explain some of this avoidance. Contempt for the common people is still widespread in many universities, though not so openly expressed as formerly. A recent interesting book about one of our oldest English universities expresses it explicitly in many places, and it would be dishonest to pretend that one does not hear it in ordinary conversation. It was even suggested recently at a scientific congress as a matter for serious study by an anthropologist.

There is, however, a subtler reason for this avoidance. I think it can be described as a belief that since a subject which interests everybody has no rarity value, he who writes about it will have no honour in his own country. For years everyone has known that millions of people, excluding ourselves, crowd in the summer to Coney Island and to Blackpool, yet a few *de-haut-en-bas* references to this weakness, as an illustration of men’s gregarious instinct, were as much as the average social psychologist felt that these phenomena deserved. Recently, an account of mass observation was broadcast in Great Britain, and the

usual comment was made (not, perhaps, with complete fairness) that the universities took no interest in it. A critic in one of our most democratic English newspapers immediately asked who could expect Oxford University researchers to be interested in what a shopgirl said on her holiday at Blackpool. Would Oxford be interested in an authentic record of what a slave girl in ancient Greece said about her mistress?

Some of the most striking omissions from general treatises on psychology seem to concern just those matters which make life interesting. Many a social psychologist seems conscientiously to have squeezed all cultural values, charm, wit and fun out of his raw material before he began work upon it. The films and the radio are two of America's chief interests, and their influence upon the success of fashions is very considerable. The radio may become the chief maker or breaker of the world's peace. Yet, so far, the number of psychologists regarding these as attractive subjects for study is very small. In some towns, England's early attempts at organizing people for air-raid precautions were made difficult by the social stratification of the inhabitants, but this subject is only beginning to appear in psychology books. Before the Great War, the idea of the inevitability of warfare seems to have been challenged by few psychologists other than William James, and to-day plenty of them seem to get along happily without bothering about such problems.

VIII. FIFTY YEARS OF CHANGE IN AMERICAN PSYCHOLOGY:

PROF. G. W. ALLPORT'S SUMMARY

It might be thought that this criticism, even if justified, is directed chiefly at the work of British psychologists, who, after all, are few in number. Yet for several years it has appeared that something similar could fairly be said of our American colleagues' activities, taken as a whole. I felt, however, reluctant to indict a whole nation, but when the above-mentioned article of my own was in proof, I found that Prof. Gordon W. Allport had done it. His presidential address on 7 September 1939, to the American Psychological Association at Berkeley, California, is reprinted⁽¹³⁾. He reminds his audience that, at a time when darkness has descended over the European continent, the burden of scientific progress in psychology rests as never before upon the membership of this Association. He continues:

With the responsibility for the preservation and eventual rehabilitation of world psychology falling upon our shoulders, we do well to examine our credentials. Are we American psychologists equipped for the versatile leadership demanded by our comprehensive discipline; are we prepared to develop the potentialities of all its parts?

Prof. Allport then asks questions which, so far as I am aware, have not appeared in the publications of psychologists, in America or this country, for a long time. I might go further, and say that in some quarters they have been studiously avoided. One possible reason for this is that until very recently, the tendency to inquire into the social desirability of the activities of any particular scientist (I do not confuse the issue by deifying and therefore placing beyond criticism that restricted kind of knowledge which is called Science) has been unfashionable, especially among the investigators of inanimate matter. One of the earliest voices raised in warning about this reckless development of physical science, with little respect for its social dangers, was that of a psychologist, the late Prof. William McDougall, in *World Chaos* (14). This message, now nine years old, was that physical science has been the principal agent in bringing about the very rapid changes in our social, economic and political conditions which are the source of our present troubles, and that our only hope of remedy for them lies in the development of the neglected social sciences. The effect of this book upon English reviewers seems to have been infinitesimal. One of our most liberal-minded English daily newspapers, noted for its reviews, gave it rather less than one-fifth the space devoted on the same page to the second volume of a book on French wines.

Prof. Allport and Mr Jerome Bruner examined all the articles in fourteen published journals of psychology between 1888 and 1938 (13). They were the *Psychological Review*; *Journal of Experimental Psychology*; *American Journal of Psychology*; *Journal of Abnormal and Social Psychology*; *Journal of General Psychology*; *Psychological Bulletin*; *Journal of Psychology*; *Pedagogical Seminary and Journal of Genetic Psychology*; *Journal of Social Psychology*; *Character and Personality*; *Journal of Educational Psychology*; *Psychometrika*; *Journal of Comparative Psychology*; *Journal of Applied Psychology*.

It is thus possible to depict fifty years of change in American psychology. Among these changes, the following are important for the present argument. There has been a distinct alteration in the manner of attacking the higher mental processes (excluding perception but including learning). There is a decline in studies based upon language behaviour, as involved in learning, reasoning, concept formation, reverie or creative thinking. In 1908 they numbered 15% of the total; in 1938 about 8%. Corresponding with this decrease is an increase (5-14%) in interest in maze learning in animals and men, conditioning experiments, and all other investigations of higher mental processes by means of non-verbal

methods. "It is clear that the distinctively human function of language has a decreasing appeal for psychologists, even when the language expression is in the form of standardized tests (cf. (15)). The only increase is in those studies of higher mental processes conducted with animal subjects or with human subjects who for the duration of the experiment are rendered totally speechless."

Are we then to conclude that for American psychologists in general the act of speaking is 'trivial'? Or have they handed over its study to the physicist, physiologist and engineer? Are *Gestalt* psychologists content with the view that the atomistic attack upon speech, being inspired by views opposed to the belief in *Gestalt*, ought therefore to be allowed to die of inanition? Perhaps not entirely unconnected with this neglect of speaking and therefore of conversation is the next fact; interest in the single case has lessened. "Articles directed towards the understanding of the individual event, based upon intensive studies of clinical cases, individual persons, or single historical events, stressing the setting of the case in its life environment", i.e. views of the single case, used to be fairly popular; in recent years they have almost been ruled out of court. "We seem now", says Allport, "relatively uninterested in what the individual case can teach, or in checking our scientific scheme against the obdurate concrete event."

There is a decline in the percentage of contributions dealing with applied psychology as well as of those concerned with social betterment. In view of the flourishing activity of the American Association for Applied Psychology, and the Society for the Psychological Study of Social Issues, there is a distinct falling off in articles applying psychology to life, or pointing it in the interests of social amelioration. The conclusion to be drawn, he thinks, "is not that our membership as a whole is less interested in the usefulness of psychology, but that a certain professional cleavage is developing. Psychologists using the fourteen journals studied are, in their writings, becoming more and more remote from living issues, and more abstract in the presentation of their subject-matter. The consulting, applied and socially minded psychologists are turning to other more specialized journals. The indication is that 'pure' and 'applied' psychology are parting ways to some extent—an event which some will deplore and others welcome."

Allport next mentions the diminution of historical surveys.

"Whether the fault lies with the authors who no longer feel respect for past work or with the editorial guillotine that decapitates articles to fit our crowded journals,"

he cannot say, "but as research accumulates in our archives it is cited less frequently in our current publications.

"Naturally, there is an increase in the employment of statistical aids and in the use of animal subjects in psychological research. Synoptic systems such as those of McDougall and Stern have given way to miniature systems, and embracing theories expounded in the grand manner have yielded to diminutive theories implemented with great precision. Immediate experience is rigidly excluded from most modern systems, and in its place surrogate operational functions are sought."

IX. DANGERS OF PSEUDO-AUSTERITY IN THE PSYCHOLOGIST'S MIND

I am anxious not to suggest by a single word that I consider any one of these special directions of psychology as useless or undesirable. I would agree with those who deprecate any immodest claims that psychology can be of use in regions where it is only just beginning. But I suggest that among psychologists in many universities, a pseudo-austerity of mind, safe for themselves but dangerous to the community, has grown up. Many psychologists are afraid of being regarded as vulgar, queer, or unsound by practitioners of the older sciences. This makes some reluctant to leave the frontiers of physics and physiology, and causes others to cling cautiously to animals. Yet the next generation, noting how few present-day psychologists were interested in unemployment, in social and national relationships, in friendship and hatred, how little they did to try to prevent or even to study war, how, fascinated with the measurement of intelligence, they took no interest in its wise direction, may consider many of the present researches childishly trivial, tragically irrelevant and deservedly unpopular.

REFERENCES

- (1) BERNAL, J. D. (1939). *The Social Functions of Science*. London: Routledge.
- (2) HOGBEN, LANCELOT (1936). *The Retreat from Reason*. Conway Memorial Lecture.
- (3) FLEXNER, ABRAHAM (1930). *Universities, American, English, German*. Oxford University Press.
- (4) PEAR, T. H. (1939). Chap. I in *The Study of Society*, ed. by F. C. Bartlett *et al.* London: Routledge.
- (5) In *Nature*, 21 October 1939.
- (6) In *Occupational Psychology* (1939), XIII, 312.
- (7) MURCHISON, CARL (1935). *Handbook of Social Psychology*. Clark University Press.
- (8) FLUGEL, J. C. (1930). *The Psychology of Clothes*. London: Hogarth Press.
- (9) BORING, E. G., LANGFELD, H. S. & WELD, H. P. (1939). *Introduction to Psychology*. London: Chapman and Hall.

- (10) In the *Brit. J. Psychol.* (1933), xxiv, 454.
- (11) PEAR, T. H. (1940). "The psychology of psychologists." *Bull. of the John Rylands Library, Manchester*, xxiv, 101.
- (12) FLUGEL, J. C. (1933). *A Hundred Years of Psychology*. London: Duckworth.
- (13) ALLPORT, G. W. (1940). "The psychologist's frame of reference." *Psychol. Bull.* xxxvii, 1.
- (14) McDUGALL, W. (1931). *World Chaos*. Ludwig Mond Lecture. London: Kegan Paul.
- (15) PEAR, T. H. (1934). "Are linguistic tests adequate?" *Brit. J. Psychol.* xxv, 55.

THE PSYCHOLOGY OF THE HISTORY AND ORGANIZATION OF ASSOCIATION FOOTBALL

By R. W. PICKFORD

Psychology Department, Glasgow University

PART II

- VI. *British prestige and the spread of the game* (pp. 128-132).
- VII. *Rugby and Association Football* (pp. 132-134).
- VIII. *Morale and the history of the game* (pp. 134-138).
- IX. *Politics and the morale of group games* (pp. 138-141).
- X. *Conclusions* (pp. 142-143).
- References* (pp. 143-144).

VI. BRITISH PRESTIGE AND THE SPREAD OF THE GAME

It is very interesting to view the expansion of Association Football as an example of the transmission of a branch of ceremonial culture (cf. Bartlett(3)). If we draw distinctions between the transmission of culture by borrowing, by grafting and by intimate contact or blending of groups of people, then the expansion of football seems to have been an example of borrowing. If football had been imposed on peoples not accustomed to play any group games, as a ceremony rendered necessary or compulsory on account of ulterior motives, by politically stronger groups, then we should view its expansion as a form of grafting. If it had expanded, for example, in Britain because the Scotch, Welsh and Irish peoples had become culturally absorbed by the English and had lost their own identities as independent national groups, then its expansion would have been due to intimate group blending. It has not, however, expanded in either of these ways. It was developed among differing groups of people, such as the several British nations, and later the European and other nations, who did not become culturally identified, and none of which imposed football as an aspect of its culture upon the others from ulterior motives. It expanded because its fame spread and because small groups of players and enthusiasts visited or went to live in distant places and carried it with them. In this way it was transmitted from Cambridge to London, from London to the North of England, to Scotland, Ireland, Wales, and then to Europe and other parts of the World. No group which took it up was forced to play football as a religious, moral or political duty to any other group. Indeed it is clear that the greatest independence was shown by the North of England, by Scottish and other

national groups. The Scots have never been willing to submit to English authority. They simply agree to play the same game. These groups all took up the game because it excited their interest and satisfied a need for group games in themselves.

Two very interesting problems arise at this point for social psychology. What was the prevailing social influence which spread the game from group to group; and why did the game not change beyond recognition in the process of transmission? In the transmission of cultural elements, ceremonial or material, there are usually great changes from group to group, and if there are no such changes it is usually important to ask why. We cannot lay down an unvarying law that cultural elements must change in transmission. We can, however, say that the degree to which such elements are likely to change or to be preserved depends on the balance of tendencies playing on those elements in their new settings. Since they always acquire a new cultural setting when they are transmitted, then they always come into the sphere of influence of a new group of active social tendencies. For example, the Christian religious culture complex has, as we can see from *The Green Pastures* (Connelly⁽⁵⁾), become modified in most interesting ways in passing into the sphere of influence of the culture of the American Negro. It has acquired certain qualities in its vividness and simplicity, and in the concrete details of its imagery, which no European or Asiatic Christian could possibly have invented by any stretch of imagination. It has been changed so much that we feel a special delight in seeing essentially the same ideas in such an astonishingly different guise. The drama of the Old Testament has been clothed in new garments quite inappropriate to Judaism, but derived from the Negro cultural context and tendencies and perfectly appropriate to these. This is a most interesting example of the transmission of a ceremonial culture complex and of the creative changes brought about in it by the play of social tendencies in its new setting.

In passing to China, the Argentine or New Zealand, however, Association Football has undergone no such dramatic changes. It has remained the same game in all essential details, and is played in these parts according to the same laws of play. The Islington Corinthians Football Club, between October 1937 and May 1938, sent a team on an extended tour of the World. This team travelled through Europe, Egypt, India, Burmah, China, Japan,* Honolulu, America and Canada, and played nearly a hundred matches, and there was no difficulty as far as play was concerned, because the same laws were adopted everywhere. Indeed, they could not have made a success of such a tour unless Association

Football had been preserved from cultural change in the course of its transmission to these widespread places.

It is certainly true that there are and always have been differences between the kinds of football played in Scotland, the North of England, the South of England and other parts, but these differences do not amount to more than divergences of style. They do not affect the essential spirit or principles of the game, and do not prevent teams from different parts playing together. The Scots players have always developed short passing and dribbling the ball, as compared with the English method of long passing and fast running. It is safe to say that no Brazilian footballer, unless trained intensively in Scottish traditions—and probably not even then—could rival one of the great Scots players in footwork. This difference, however, would not make it difficult for a Scots and a Brazilian team to enjoy a match together. A Scottish player would not have the same degree or kind of difficulty in playing in a Brazilian team that an English bishop would have in trying to conduct a Negro Christian religious service.

The essential reasons why Association Football has spread so widely with so little change are simply that English and British prestige were strong influences in its cultural transmission, and the pioneers are still felt to understand the game best; and that the essential condition for the success of a group game like football is co-operative and competitive enterprise based on fair play. These two factors have ensured that the game should be passed on and at the same time preserved in its essential form according to the Laws and Rules of The Football Association. Of course, in Africa the natives may play barefoot, but that is only an adaptation suiting the game to people who have strong and capable feet. Groups which took up the game in the course of its wide expansion did so because of its reputation as a great game spread abroad by the British, who were pioneers in group games. These new groups tended to accept the supremacy of the British in athletics⁽¹⁹⁾, and, since they were motivated by the desire for co-operative competition, they made no moves to assert their independence by alterations to the essential laws of the game.

These points are all of great interest because, in general, prestige alone does not seem to be able to preserve travelling elements of culture unchanged. If it was sufficient alone, then many practices and customs which have been introduced to native peoples all over the world by Europeans should have remained more constant than they have. In all likelihood transmission without change is more frequent the less there

132 *History and organization of Association Football*

are already in the cultures to which the elements go, settings closely related to the incoming materials or customs. In Association Football an important additional factor seems to have been the need for widespread common understanding about the comradely and competitive enterprise. It is obvious that religious practices would be expected to undergo great change, in view of these arguments, because they usually meet with pre-existing religious practices in their new settings, and their transmission is often in part the result of active conflict with parent groups. British group games, however, seem to have had an open field in which to spread abroad.

VII. RUGBY AND ASSOCIATION FOOTBALL

Handling the ball has become the great distinction of organized Rugby Football, while in the Association game it is strictly taboo for the ball to be touched by the hand at all when in play, except by the goalkeeper. It is, however, thrown into play again when it has crossed the touch-line. The concentrations of interest on holding the ball and running with it, in Rugby, and on playing it only with the head or feet, in the Association game, are of great psychological interest. It might be said that in Association Football the ball is symbolically a very dangerous or dirty object, or that it is so powerful in unconscious meaning that it must not be touched with the hand under penalty for foul play. In Rugby Football the ball might be regarded as such a lovable object that every player tries to caress and hold it as if it were a valued possession. In psycho-analytic terms these attitudes suggest that, to the Rugby enthusiast, the ball is a symbol of loved maternal images, while to the Association Football enthusiast it is a symbol of paternal potency which is somewhat feared and requires special forms of taboo or control. These differences are born out by the importance in the Association game of control by laws and rules and on the hierarchical system of organization, which is, nevertheless, very democratic. In the Rugby game the ball is put upwards over the goal bar, while in Association Football it is kicked beneath the arch. These differences may carry the symbolisms of the ball further, because 'upwards' suggests purity and 'downwards' suggests the mastery of a rival. Shooting a goal in Association Football may also have a sexual meaning. On the other hand, the ball is kicked, headed and passed from player to player with loving care and skilful footwork in Association Football, and in the Rugby game the 'scrum' might be regarded as an ugly form of aggression and even as an undignified tussle for possession of the loved object, the ball. It is, therefore, easy for

partisans of either game to represent the other as objectionable and inferior, but in fact each has its merits as a game. The two games provide opportunities for the gratification of very different tendencies in individual players and spectators—Association Football gratifies an interest in skilful footwork with a special taboo on touching with the hand, Rugby in handling the ball, and so on. The growth of divergence between the two kinds of game expresses the influence of strong social class distinctions, and it is an excellent example of what Bateson (2) has called 'schismogenesis'—the development of cultural traits in opposition and divergence instead of towards blending and unity. The divergence here was based upon class distinctions together with important unconscious meanings in handling the ball as opposed to footwork with a taboo on touching.¹

At the meeting in 1863, when the supporters of the handling and carrying game withdrew, and left the formation of The Football Association to the men interested in the kicking and dribbling game, the representatives of handling had been instructed to insist upon hacking as an essential part of football. The insistence on hacking, which they thought to be essential, was really an unconscious excuse for the great division between the supporters of handling and those who upheld the taboo on touching the ball. It is an interesting example on the social level, of the way in which people are commonly found to act as they wish, but to insist on reasons which are not their real motives, which are unconscious. In 1870, when the Rugby Union was formed, hacking was given up in the Rugby game too. At first the divergence between the two kinds of football was not dependent on social class distinctions, but simply on traditions of school-play and on preferences for handling and objections to it. Later, Association Football tended more and more to become the game of the people, and Rugby the game of the upper middle classes and the English type of public school. Although both games emerged from the traditional background of medieval football, they have become sharply contrasted, and the contrast is of much social and individual significance. It is not comparable, however, with the distinction between two different games, like golf and cricket, and it has a far greater social and emotional importance. It is probably a general principle of social psychology that where two cultural traits, like Association and Rugby

¹ This paper was read at a meeting of the Scottish Branch of the British Psychological Society, 3 December 1938. The importance of the difference in meaning between handling and kicking with a taboo on touching the ball was stressed in the discussion at that meeting, and was specially emphasized by Dr E. H. Connell.

134 *History and organization of Association Football*

Football, have recently diverged for emotional and social reasons from a common background, the opposition between their respective supporters is usually more acute than where two cultural traits are much less alike in superficial appearance and in traditional background. Religions which are diverging from a common basis are often bitterly opposed, while religions which have never been connected may exist side by side without conflict.

It is evident that the division of interest over Association and Rugby Football has been too great, from the start, to be overcome by the impulses which have made each game, and particularly the Association game, very uniform. While most of this article is about the tendencies which made Association Football a widespread and uniformly organized system, and which overcame threatened splits, it is very interesting to contrast these influences with the profound likings for handling the ball and for the taboo on touching, which were the basis for the divergence of cultural growth between the Rugby and Association games.

VIII. *MORALE AND THE HISTORY OF THE GAME*

The early forms of football seem to have been ceremonial combats, and it would be very interesting to investigate the unconscious meanings of the ball itself, which was, in some forms of the game, possibly a symbol of paternal authority and potency (see Elliot Smith⁽⁷⁾ and Perry⁽¹⁵⁾). The early and medieval games were violent, and many of the players in them received serious injuries. In the nineteenth century the tendency was towards less violent games, and, when the Cambridge Rules were put together, hacking was given up for those who followed these rules, and a relatively gentle form of football established, in which skill took the place of violence. With the development of six-man football in the U.S.A., since 1934, a similar change towards the replacement of violence by skill seems to be coming in American football. "Regulation football's lasting glory has been due to the fact that it is basically brutal and elemental... The six-man game is the very opposite of regulation football..." ("Mr Football Becomes a Gentleman"⁽⁶⁾). Although the men who established the Rugby game broke away in 1863 in part because they liked hacking, this has not survived in Rugby Football, though certain other somewhat violent aspects of the early games have been retained in it.

Many incidents in the history of the Association game show that the impulses of fair play and of mutual consideration were important, not

only on the field of play, but also in the council room. The capacity of the Football Association for compromise over the demands of professionalism, and over the problems of the Amateur Football Alliance, are illustrations. The freedom of the Scottish, Irish and Welsh Associations, and of the F.I.F.A., to control their own internal affairs, while they still adhere to the laws of play approved by the International Board, and their representation on the Board on an equal footing with The Football Association, though some might be regarded as more important than others, are further illustrations of the influence of impulses of comradeship in the control of the game. An obvious example of the influence of *morale*,¹ as contrasted with discipline, is the unwillingness of the F.I.F.A. to lose harmonious contact with the International Board, and of that Board with the F.I.F.A., in spite of differences of opinion about the definition of professional status of players in Britain and on the Continent. No rational or moral intentions of individuals, nor any dictatorial law-making, could have set up and maintained the existing harmonious organization of the game.

The Cambridge, London and F.A. rules and laws were established and adopted because many players needed some simple code which they could easily grasp and which would make fair play practicable and enable them to enjoy games together. They wished essentially for a practical understanding and to avoid quarrels and differences, both on and off the field, which might spoil the game. The Football Associations and the great organizations of League Football arose from the same essential needs, and the Associations and Leagues have their authority simply because football is a social institution in which a very large number of people want to enjoy their parts, either as players or as spectators. It is not surprising that people who did not accept and fall in with these principles have, on the whole, either had to keep quiet or to withdraw.

The essential principles can be summarized as follows: that football should be conducted for the advantage of the public, players and spectators, and not merely for private fame or gain; that where difficulties could not be resolved directly, it was better to compromise than to allow misunderstandings to interfere with the value of the game for the public; that no individual, club or association taking part in the game could be omnipotent or have the right to interfere with the private affairs of any other party. These principles all arose out of the essential comradesly qualities of group games, and it is most unlikely that such principles

¹ Cf. Bartlett (4) and Reaney (17). Patrick (14) advocated the recapitulation theory of play for American football.

could have been applied in the control of international football unless they had been essential to the game from the very start.

A further aspect of group games, which is of the greatest psychological importance, though it could not be called a principle, is good humour. The social significance of good humour appears again and again in football, not only on the field of play itself, but among the spectators, and especially among the members of the organizing bodies, such as the International Board. Anybody who has joined the crowd at an International match, for example, Scotland *v.* England at Wembley, where national feeling rises to a high pitch of excitement, must realize that organizations, rules and laws would all be useless to achieve fair play unless the crowd took part in the whole proceedings with good humour. Again, the International Board might easily find its problems insoluble unless good humour were traditional at its meetings. This capacity for good humour makes social co-operation possible under the most trying conditions, and prevents the essential comradeship from vanishing at moments of stress.

Many people do not realize the futility of laws and rules in themselves to control a game like football, or human conduct in general. Laws and principles always have to be backed by some power derived from the impulses of human nature itself in order to be effective. In political organizations they are only too often backed by impulses of fear owing to threats of violence and punishment. In most problems of law and administration it is necessary to have possibilities of punishment in the last resort. This is very unfortunate. The extent to which a controlling institution is able to count on public support for its acts, without resort to punishment, measures the degree to which its members are controlled by *morale* rather than by discipline. Punishment is essential to discipline; fair play and good humour are the foundations of *morale*. The kind of flexible and democratic organization maintained by The Football Association would have been impossible on the basis of discipline and punishment, and it is in the organized group games that *morale* is most adequately developed.

The expansion of Association Football discussed in this article took place through the interaction of group and public interests with the influence of gifted men who were great leaders, through the intermediation of small groups consisting of clubs, associations and controlling bodies. It is difficult to discuss this very interesting three-cornered system adequately without analysis of the activities of well-known men some of whom are still living. It is commonly found that the relationships of

men of genius, small controlling groups and the wider public setting are the foundations of social systems in which there is expansion and constructive change of culture. A great interest for social psychology in the history of Association Football is the dependence of the game on *morale* and impulses of fair play. Owing to these factors its history is a contrast to the history of religions, of political and artistic cultural movements, which usually show endless schisms and dissensions. The history of the game during the period discussed was under the influence of such men as Lord Kinnaird, Sir Charles Clegg, D. B. Woolfall, and the leaders who preceded, followed and worked with them. A new generation of leaders and the social changes of the mid-nineteenth century may bring unexpected developments in the game and its organization. The enthusiasm of the public for the 'Football Pools' may change the attitude of many people to the game, though the interest in betting is strongly disapproved by The Football Association. The war may bring many changes which we cannot at present foresee.

Outside the British Isles international football is taken much more seriously than the British tend to take it. However warm national feeling may be about international matches between the countries of the British Isles, it seldom reaches such a pitch of seriousness and intensity that the players and spectators are more interested in winning than they are in the game. When British teams go to play abroad it is often in the summer; then they are having what might be called a holiday, and they do not mind so much about winning as do their Continental or other opponents. Foreign enthusiasts find this very difficult to understand, and there are occasions on which even the all-important good humour vanishes. This is well illustrated by the story of the South American spectator, who is reported to have become so excited and angry that his side was losing, that he whipped out his revolver and shot the ball. Although international football is a most interesting sublimation of crude aggression, it has limited capacities in this direction. It has not proved itself a substitute for war, and in some European countries it is almost tending to become a direct expression rather than a sublimation of international rivalry. Then it is being taken too seriously.

If the principles of international football mentioned above could become traditional in wider spheres, many difficulties between peoples and nations might more easily be solved or avoided. The sort of *morale* developed in connexion with group games, which can maintain an international organization on the basis of fair play for everybody, compromise in place of aggressive insistence on differences of principle

138 *History and organization of Association Football*

and the avoidance of interference with other people's affairs, is very important indeed.

The transmission of Association Football abroad with the minimum amount of friction and adherence to good principles of co-operative enterprise, is closely related to the democratic control of the game. This control became firmly established before the organization involved spread widely in foreign countries, and it therefore had prestige and the advantage of time in its favour. It is particularly interesting that the democratic control has not changed although the game has spread to countries the general character of whose political organization is autocratic or despotic. This may mean that democratic control is not impossible anywhere in the world on account of anything inherent in the method itself. Where it is superseded as a method of social rule by something more despotic, we should perhaps look for the reasons not inside the method of control itself, but rather in the spheres of culture concerned. Further, compromise, with occasional drastic action, does seem to be inherent in this form of control, but in the sphere of group games it does not wreck the method, whereas in other spheres it has done so. If this game was transmitted with lack of friction because The Football Association had already a very firmly established prestige before much spreading took place, then it forms an excellent illustration of the functions of social prestige in facilitating agreements between group and group.¹

IX. POLITICS AND THE *MORALE* OF GROUP GAMES

In view of these points it is interesting to consider some further aspects of international football. For example, from the point of view of social psychology, it is very important to try to understand why international football has, on the whole, tended to be a sublimation of aggression rather than a direct expression of group conflict. Throughout their long history, from their ancient origins to the present day, football and the group ball games from which it arose have tended to be sublimations of aggression. They have taken up the social functions of substitutes for injurious and destructive conflict. They have been symbolic rather than actual fights. Medieval monarchs tried to suppress them because they deflected interest from war-making. How much better it would have been to have suppressed war because it interfered with football! Only where the essential principles of compromise, of good humour, of

¹ For advice and criticism about this and other parts of the discussion I am much indebted to Professor F. C. Bartlett.

non-interference with other peoples' private affairs, and, in general, of fair play, have broken down, has football tended, as between certain European countries recently, to become the active expression of aggression.

There are important and puzzling problems for social psychology, about the reasons for the breakdown of these principles. War often arises on account of precipitating causes (whatever may be its fundamental causes) which seem to depend on a failure of the same principles. It seems to be caused by failure of the parties concerned in a dispute to be able to compromise, by their tendency to interfere with each other's private affairs, by lack of sense of humour, or, in general, by incapacity for fair play. It is usual to appeal to reason when these principles break down, and to ask why one or both parties to the dispute cannot be moral and reasonable. No amount of appeal to reason and to public or international morals, however, will produce reasonable and moral attitudes by itself, where no tendencies for fair play and no sense of humour exist, or where they cannot be created. This is one of the great problems of psychology. The effect of appeal for moral and reasonable action depends not on reason itself, but on the fundamental psychological relations holding between the group or individual who makes the appeal and the individual or group which receives it. If the relationship is fundamentally one of antagonism, then appeal to reason may result in behaviour more unreasonable than ever. Reason and morality only seem attractive when they accord with fundamental motives, and some of these, in any personal or group relationship, are usually unconscious. This does not mean that reason and morality are only justifiable in the abstract when they seem attractive in the concrete; it means simply that they are not themselves motives, but are dependent on the deeper sources of human motivation.

International football has been fortunate. The success and interest of the game depends entirely on fair play and good humour. The public and players have usually been so interested in the game for the game's sake, that fair play and good humour have generally been possible in order that the game should have continued interest and success. Only under these conditions can it be a good sublimation of international aggression. There remain profound problems, however, about international politics, the battle-ground for innumerable ulterior motives, and in which no party is interested 'for the game's sake'. How could the methods of politics be modified to secure more sublimations of conflict than in the past, and to lead less often towards direct expressions of international aggression?

It is shown by the history of the game, that an important aspect of the development of Association Football, both national and international, has been the establishment of a tradition of reasonable co-operation in the place of downright aggression in the settlement of disputes and difficulties. Human social life is controlled by tradition to a remarkable degree. Conservative tendencies, which we may suppose to have the primary function of stabilizing social systems and cultural patterns, lead towards the formation of traditions which operate strongly in most spheres of social activity. These traditions, and their relations to social stability and to social change, are a very large subject of study, and they cannot be analysed in any detail here. It is possible, however, to say that the tendency to follow tradition is very strong in mankind, even though often opposed successfully by rebellious tendencies. It would be a grave breach of tradition in the organization of Association Football for any problem relating to the game to be settled by violence rather than by peaceful means. In fact, precisely because such use of violence would be a great breach of tradition, it appears an absurd possibility that any but a peaceful method should ever be adopted. We should be extremely surprised to hear that The Football Association Council had been forced to agree to a change of which they disapproved in the Laws of the Game, under threat of being shot if they refused. It would be equally amazing if, for example, the referee at an International match between England and Scotland should be forced at the point of the bayonet to grant a winning goal to either side. Such astounding breaches of tradition would be strongly opposed. The public would not take a rational view and give such possibilities their serious consideration. It would not be questioned whether the adoption of these violent methods might bring unfortunate changes in the constitution of the Association, and in the kind of organization which is customary in football. It would not even be questioned whether the public might possibly approve of such methods. The public interested in football would certainly reject them straightway, without giving them any consideration whatever. Opposition to such methods, or even to the threat of their use, would be made solidly by all supporters of the game. The possibility of such methods can be expressed to us intelligibly only in the form of a joke (see "The Ten Little Referees," Pickford (16) p. 100).

In contrast with this, it is worth noting that potential threats of resort to violence are always in the background in international politics. Sometimes such potential threats materialize into serious warnings, and sometimes they become real threats of war. Only too often it has been

necessary to resort to war or threat of war in order to secure peace. In international politics, we might say that it is traditional to resort to war as the ultimate and supposedly satisfactory method of solution of major problems, and that war has been traditionally glorified as an ultimate measure in the settlement of disputes. The future hopes of world peace seem to depend at least in part on the capacity of the peoples concerned to change this tradition of ultimate resort to war. The Football Association never had to face such a problem, because the game was relatively unorganized while it was traditionally violent, and the Association, the International Board and the F.I.F.A. came into existence because of the rising public interest in the co-operative and peaceful organization of the game. If an international organization could come into existence equally spontaneously with the development of the Football Associations, because of a rising public interest in the peaceful settlement of international disputes and problems, and could command the support of all the great powers concerned, then war might become the method of settlement to be traditionally avoided in the future.

It is not, however, likely that such an international organization could ever be created with success artificially, or as a result of the defeat of its opponents in a previous war, or unless it could win the whole-hearted support of all the important national powers concerned.

In connexion with this problem, we must observe that it is often extremely difficult for an arrangement which has proved quite workable in one sphere of culture to be transferred to another. This may be a particular form of the very difficult problem of the transfer of training, appearing in a social setting. It is a question whether methods of social organization and control which have been evolved for football could be transferred to political and other spheres; whether, if people acquire habits of co-operation in group games, they will be made more co-operative in politics. The real difficulties in such a transfer are probably emotional, and are not concerned purely with the psychology of motor and cognitive learning. The emotional meanings of group games, which lean predominantly towards the sublimation and limitation of aggressive impulses by *morale* and fair play, may contrast too sharply with the chief meanings of political interests, which have been predominantly aggressive in the past, for much transfer of training to take place. The traditional British interest in group games, and their traditional ability as colonists and organizers, may spring from the same fundamental source.

X. CONCLUSIONS

A brief account has been given of the development of organized Association Football in England, and of its expansion in Great Britain, Europe and other parts of the World up to the outbreak of war. Many changes may take place when football is resumed in its normal way after the war is over. The impulses of fair play seem to have had tremendous importance in the development of organized Association Football. The change from the violent games, which were its forerunners, to regularized football, was essentially due to the growing public desire for harmoniously organized group games of interest to spectators as well as to the teams and their immediate supporters. Group games such as football are symbolic rather than actual combats between teams whose 'sides' are represented by the spectators who enjoy the game with sympathetic interest. A very high degree of individual skill is called for in such games, together with an equally high degree of co-operative team work. The team work is by no means confined to the individual teams taken singly; though each of them wishes to win the game, and they are not on that account merely opponents. The team work involves both sides in a double team whose object is to have a good game. Association Football also calls for a high degree of public organization, of which the purpose is to maintain the standard of play, to ensure that the teams which meet in the field are well matched, and to arrange for co-operation between all the many clubs and associations and between the national groups interested in promoting football.

It is of great interest that Association Football has spread widely by a process of cultural borrowing from group to group. In this it has undergone surprisingly little change, as compared, for instance, with the changes religious and artistic culture complexes usually undergo in the course of wide transmission. This lack of change is probably due to the predominance of British prestige in its transmission, and to the prominence of the impulses towards fair play and the *morale* of group games. The main desire in international spread of the game has been for co-operative competition rather than for direct conflict and the consequent separation and cultural divergence of groups.

The histories of The Football Association, the International Board and the F.I.F.A. show that the fundamental nature of the public interest in the game has enabled these bodies to conduct their affairs, and especially their international affairs, without resorting to threats, violent measures, and schisms. They have contrived to meet their most difficult

international problems in a spirit of compromise, with humour, without interfering with each others' private affairs, and in such a way that the public was considered before private gain and personal glorification. These principles have been maintained because men were inspired by the fundamental co-operative nature of the group game, and not because they set out deliberately to try to be moral. The result has been the establishment of traditions of fair play and co-operative methods of international organization, such that any other methods or principles would now be solidly opposed by the public as absurd because contrary to tradition.

It is of great interest to consider how international politics might be made to follow the same principles and how the same traditions of the avoidance of the threat or actual use of violence might be established in them. This might be possible if the great international powers could spontaneously become interested in co-operation for the benefit of mankind rather than in the glory and success of one party at the expense of others. It is essential that such co-operation should be spontaneous, and not forced upon a loser by the victory of his opponent. It is not likely that such co-operation will be achieved by the efforts of deliberate pacifists, though they might contribute to the necessary change in political traditions.

REFERENCES

- (1) ALVERDES, F. (1927). *Social Life in the Animal World*, pp. 160-3. London.
- (2) BATESON, G. (1936). *Naven*, Ch. XIII. Cambridge University Press.
- (3) BARTLETT, F. C. (1923). *Psychology and Primitive Culture*, Chs. v-viii. Cambridge University Press.
- (4) — (1927). *Psychology and the Soldier*, Ch. vii. Cambridge University Press.
- (5) CONNELLY, MARC (1929). *The Green Pastures*. New York.
- (6) *Digest and Review* (1939). vii, no. 1, 94-5. New York.
- (7) ELLIOT SMITH, G. (1933). *Human History*, pp. 304-12. London.
- (8) *Fédération Internationale de Football Association*, Handbook (1938).
- (9) GARDINER, E. N. (1930). *Athletics of the Ancient World*, Chs. II and XVIII. Oxford University Press.
- (10) GIBSON, A. & PICKFORD, W. (1906). *Association Football*, vol. i. London.
- (11) JACKSON, N. L. (1900). *Association Football*. London.
- (12) LANGENUS, J. (1938). "From Cambridge to London." *The Football World*, no. 1, 12.
- (13) *London Illustrated News*, 27 Aug. 1938, pp. 353-6.
- (14) PATRICK, G. W. (1903). "Psychology of football." *Amer. J. Psychol.* xiv, 104-17.
- (15) PERRY, W. J. (1923). *The Children of the Sun*, pp. 312-13. London.
- (16) PICKFORD, W. (1939). *A Few Recollections of Sport*. Printed by the Bournemouth Guardian, Ltd. (Private circulation.)

144 *History and organization of Association Football*

- (17) REANEY, M. J. (1916). "The psychology of the organized group game." *Brit. J. Psychol., Monog. Supp.* no. 4.
- (18) RICKMAN, JOHN (1938). "A discursive review: Air raid." *Brit. J. Med. Psychol.* xvii, 361-73.
- (19) RIMET, JULES (1938). "A symbolic match." *The Football World*, no. 1, 1 and 2.
- (20) *Rules of the Association and Laws of the Game*. The Football Association, Season 1937-1938 (handbook).
- (21) SHEARMAN, M. & VINCENT, J. E. (1872?). *Foot-Ball: Its History for Five Centuries*. London.
- (22) SUTTIE, I. D. (1935). *The Origins of Love and Hate*. London.
- (23) YAMANO (1938). "Le Football au Japon." *The Football World*, no. 1, 18.
- (24) FAIRBAIRN, W. R. D. (1938). "Prolegomena to a psychology of art." *Brit. J. Psychol.* xxviii, 288-303.
- (25) LEES, DOROTHY N. (1939). "Pageants of Tuscany." *Nat. Geog. Magazine*, ix, 191-212.

FURTHER DEVELOPMENTS IN A NEW APPROACH TO THE TESTING OF MUSICAL ABILITY, WITH SPECIAL REFERENCE TO GROUPS OF SECONDARY SCHOOL CHILDREN¹

By BORIS SEMEONOFF

Department of Psychology, University of Edinburgh

- I. *Introduction: scope of the experiment* (pp. 145-146).
- II. *Testing programme* (pp. 146-151).
 - (1) *Test of musical understanding* (pp. 146-147).
 - (2) *Test of musical taste* (pp. 147-149).
 - (3) *Test of musical knowledge* (pp. 149-150).
 - (4) *School attainment* (p. 150).
 - (5) *Administration of the tests* (pp. 150-151).
- III. *Test results* (pp. 151-156).
 - (1) *Test of understanding* (pp. 152-153).
 - (2) *Test of taste* (pp. 153-154).
 - (3) *Test of musical knowledge* (p. 154).
 - (4) *Correlations between tests* (pp. 154-155).
 - (5) *Sex differences* (pp. 155-156).
- IV. *Discussion* (pp. 156-158).
 - (1) *Test of understanding: responses of children of different ages* (pp. 156-157).
 - (2) *Test of taste: validity of instructions and of general method* (pp. 157-158).
 - (3) *Suitability of technique for the groups tested* (p. 158).
- V. *General conclusions* (pp. 158-160).
- VI. *Summary* (pp. 160-161).
References (p. 161).

I. INTRODUCTION: SCOPE OF THE EXPERIMENT

THIS paper may be regarded as a continuation of an investigation reported in a previous issue of this *Journal*⁽⁴⁾. The present study was undertaken with two main purposes in view: (i) to extend the method of music testing already tried with adults to secondary school children, and (ii) to devise, if possible, a test of musical taste on similar lines to those of the test of musical understanding which formed the subject of the earlier investigation. From the results it was hoped it might be possible to draw conclusions regarding the existence or otherwise of an innate

¹ Based on a paper prepared for delivery at the Dundee meeting of the British Association for the Advancement of Science, 1939.

unitary musical factor or ability, as distinct both from a group of more specific abilities, and from proficiency or knowledge attained as a result of musical education.

The choice of subjects for the experiment was determined partly by convenience, and partly by the conditions of the previous experiment. Since the subjects in that case were nearly all University students, it was decided to take as closely similar a group as possible, and to confine the investigation to age groups within the secondary schools proper, so that the method and material would require the minimum of alteration. Permission was obtained from the heads of three schools to test top divisions of the first three years of the secondary department, these children being regarded as typical of those most likely to embark later on University studies. The schools chosen were fairly representative of the educational background of Edinburgh University students—one boys' school, one girls' school, and one mixed school under the auspices of the Education Committee. The four groups thus formed are referred to in what follows as B (boys), G (girls), mB (mixed school boys), and mG (mixed school girls). In all, 270 subjects took part, ranging in age from 11 years 7 months to 16 years 7 months, but for practical purposes the group consisted of children between the ages of 12+ and 15+ inclusive.

II. TESTING PROGRAMME

The testing programme comprised the following: (1) and (2) tests, as already stated, of musical understanding and of musical taste, in which impressions received on listening to gramophone records were noted on specially prepared blanks; (3) a simplified version of the 'Revised Kwalwasser' test described in the previous paper, entitled "George Combe Test of Musical Knowledge (J)"; (4) in the case of one school, a list of marks awarded by the music teacher for the work which the children had done for her was obtained.

A detailed account of the test material follows.

(1) *Test of musical understanding*

The nature of this test will be clear from the following instructions, which were printed at the head of the test blanks:

You will be asked to listen to a series of gramophone records. After each has been played, answer the question printed opposite the number of the record, by underlining the word or phrase which you think gives the best description of the music. Explanations of some of the more difficult words are given on the page of notes, under the number of the record. You are also asked to say whether you liked each record or not. The answer should be put on the dotted line after the number of the record.

The records used were selected from those of the previous experiment, those records being chosen which appeared to be most suitable for the juvenile group. The following is a list of the records used in this instance:

1. (A2) Lieurance (arr.). *By the waters of Minnetonka*. (Cherñiavsky Trio.) Col. 3368.
- 2.† (A3) *Volk ans Gewehr*.* (Band and Male Chorus.) HMV. EG 2775.
3. (B4) Chaminade. *Sérénade Espagnole*. (Violin: Bratza.) Col. 5466.
4. (B6) Saint-Saëns. *Carnaval des Animaux: Tortues*. Col. 9519.
5. (A4) Saint-Saëns. *Carnaval des Animaux: L'Éléphant*. Col. 9519.
- 6.†. (A1) Mozart. Minuet from *Don Giovanni*.* (Harpsichord and Orchestra.) HMV. DA 977.
7. (A5) Sibelius. *Finlandia*. (Symphony Orchestra.) HMV. DB 1584.
8. (B2) Briccialdi. *Il Vento*. (Flute: Amadio.) Aeol.-Voc. R. 6024.
9. (B1) Knipper. *Song of the Plains*. (Choir of the Red Army of the U.S.S.R.) Col. DB 1749.
- 10.†. (A9) Debussy. *La Fille aux cheveux de lin*. (Violin: Heifetz.) HMV. DB 1246.

Numbers in brackets indicate the series and serial number in the previous experiment. In two cases (marked *) less than the whole record was played.

The page of notes contained easy explanations of every word or phrase which might conceivably cause difficulty. This expedient was adopted to avoid possible loss of time through answering questions. Some slight changes were also made in the form of question as printed on the test blanks, but in general a policy of minimum change was adopted. Only in the cases of selections marked (†) was any change made in the actual alternatives offered, and these changes, too, were slight. As a result, a direct comparison was possible between scores made by school children and those obtained on the same items by the adult groups.

(2) *Test of musical taste*

The method used for this test was based on a technique described by Trabue(5) in a report of a little-known experiment by M. L. Mohler. The test blanks used bore the following heading:

Gramophone records will be played in groups of three. In each group, the first record played is called A, the second B, and the third C. After each group of three has been played, decide which you liked best, which next best, and which you liked least. Then put a figure 1 in the column headed by the letter corresponding to the record you liked best, a figure 2 in the column corresponding to the record you liked next best, and a figure 3 in the column corresponding to the record you liked least.

In the Mohler-Trabue experiment only recordings of orchestral music were used, but for a variety of reasons it was thought preferable to

include piano and violin works. Accordingly, three groups, each consisting of three records, were played: Group I, orchestral records; Group II, piano records; Group III, violin records. Unfortunately, it was not found possible to determine the relative merits of the works in each group by as extensive and rigorous a procedure as that described by Trabue, but the records were chosen so that each group should contain one acknowledged masterpiece of its own kind, one definitely ephemeral piece, and one of intermediate merit. The further precaution was taken of choosing as 'second-best' items, in cases of doubt, works which would not make so direct an appeal to the non-expert listener as would those graded as 'best'. It was considered that this device would have the same effect as the Mohler-Trabue procedure of averaging the judgements of 'expert' and 'non-expert' subjects.

The following is a list of the records used. Numbers in brackets indicate placings for merit (1, 'best'; 2, 'next best'; 3, 'worst'). Only the second sides of the overtures in Group I were played.

Group I. Orchestral.

- A. Mozart. *Magic Flute* Overture. Parl. E 10750. (1)
- B. Romberg. *Viennese Nights: Introduction and I bring a Love Song*. Col. DX 208. (3)
- C. Boieldieu. *Caliph of Baghdad* Overture. HMV. B 3482. (2)

Group II. Pianoforte.

- A. Grieg. *Feuille d'Album*. (de Greef.) HMV. D 1825. (2)
- B. Chopin. *Waltz in D♭ major*, Op. 70, no. 1. (de Pachmann.) HMV. DA 761. (1)
- C. Young & Petkere. *Dancing Butterfly*. (Raie da Costa.) HMV. B 4477. (3)

Group III. Violin.

- A. Kreisler. *Caprice Viennois*. (Kreisler.) HMV. DB 1091. (3)
- B. Brahms. *Valse*, Op. 39, no. 15. (Spiwakowsky.) Parl. 1414. (1)
- C. Tchaikovsky. *Melodie*, Op. 42, no. 3. (Elman.) HMV. DA 1143. (2)

The ratings were checked by carrying out on a small scale the standardization procedure described by Trabue. The judgements of a group of eighteen adults and four juveniles of over 16 years of age were averaged, and relative values of the records in each group were calculated from the formula $(x-y)/2n$, where x = number of times a given record was ranked above either of the others in its group, y = number of times it was ranked below either of the others, and $n = (x+y)/2$ = number of persons participating.

Values thus obtained vary between +1 and -1. The values are given in Table I, and it will be seen that they agree completely with the ratings originally assigned.

Table I. *Relative values of records in test of musical taste*

	A	B	C
Group I	+0.32	-0.36	+0.05
Group II	+0.18	+0.45	-0.64
Group III	-0.45	+0.64	-0.18

Scoring of this test was carried out according to the system advocated by Trabue. Two points each were given for placing the 'best' and 'worst' records correctly, and one point for placing either of these in the mid-position. This gives a possible score of 4 for each group, and no possibility for scoring 2 on a single group. Scores obtained by this method can also be interpreted in other ways, and the method appears to be logical whichever way one considers it. Chance expectation of total scores for three groups gives the following distribution, which will be seen to be approximately normal, except for a drop in respect of the mid-category:

Score	0	1	2	3	4	5	6	7	8	9	10	11	12
Proportionate frequency according to chance	1	6	12	14	27	36	24	36	27	14	12	6	1

(3) *Test of musical knowledge*

As in the case of the experiment with adults, it was considered desirable to obtain a measure of the musical status of the subjects—of how much they knew about music generally. Since the writer's revision of the Kwalwasser test was too difficult for school children, a new test was devised. This consisted of five sub-tests, containing in all twenty-eight questions, to some of which two points were allocated. Table II briefly analyses this new test, together with the 'Revised Kwalwasser' and the original Kwalwasser test(1).

Table II. *Contents of tests of musical knowledge*

Test	...	GC(J)		RK		Kw	
Type of question		Test no.	Possible score	Test no.	Possible score	Test no.	Possible score
Nationality of composers		I	5	I	10	2	20
Classification of composers		II	5	II	5	4	10
General knowledge of composers and compositions		III	8	III, IV	27	3, 5	80
Classification of instruments		IV	8	V, VI	11	6, 7	20
General knowledge of musical theory, etc.		V	9	VII, VIII	27	8, 9	100
Classification of artistes		—	—	—	—	1	20
Total			35		80		250

Of the five sub-tests, only one (test III) involved scoring on the (Right—Wrong) principle. All the sub-tests, although modelled for the

150 *A new approach to the testing of musical ability*

most part on the Revised Kwalwasser, consisted almost entirely of new items. In test V, in particular, considerable variety in the type of response was introduced, thus making the test more attractive to children.

Intercorrelations of the sub-tests, based on 210 cases, and thus all with a nearly uniform probable error of about 0.04, were as in Table III. These, it will be seen, are all low, positive, and (all but one) significant, as was to be desired.

Table III. *Test of musical knowledge: intercorrelations of sub-tests*

Test	II	III	IV	V
I	0.43	0.27	0.18	0.14
II	—	0.35	0.20	0.33
III		—	0.11	0.23
IV			—	0.26

(4) *School attainment*

Correlations were calculated between the various measures and school attainment in music, but since not all the groups could be studied in this way, and since school marks are both highly unreliable and difficult to interpret, no account has been taken of these in the present report. It may, however, be of interest to note that the obtained correlations were of a very similar order to the others found (mainly low and positive), and seemed to throw little light on the problems in hand.

(5) *Administration of the tests*

The auditory tests were so drawn up as to make each capable of administration within a single school period. In the case of the mixed school the knowledge test was separately administered under the supervision of the music teacher. In the other two cases, the papers for this test were distributed at the beginning of the period devoted to the taste test, and the children told that they could fill it up during pauses between records, or while the music was being played. It was considered that distraction of attention which might arise would be slight, and that in any case it would be quite a good thing to give the children something to do should their attention wander.¹

It was hoped to administer the two tests with about a week between them, and to take each class separately. Exigencies of time, however,

¹ It was noticed that the one class which did not do the knowledge test showed a marked openness to distraction during the period in question.

made it impossible to secure uniformity of test conditions, and the following arrangement had to be adopted:

Boys' school: Classes taken separately; 12 days' interval.

Girls' school: All classes together; successive periods with 15 min. break between—an unsatisfactory arrangement.

Mixed school: All classes together; successive days.

In all cases the test of understanding preceded that of taste. How far the conditions affected the reliability of the results is of course uncertain; probably the effect was not very great.

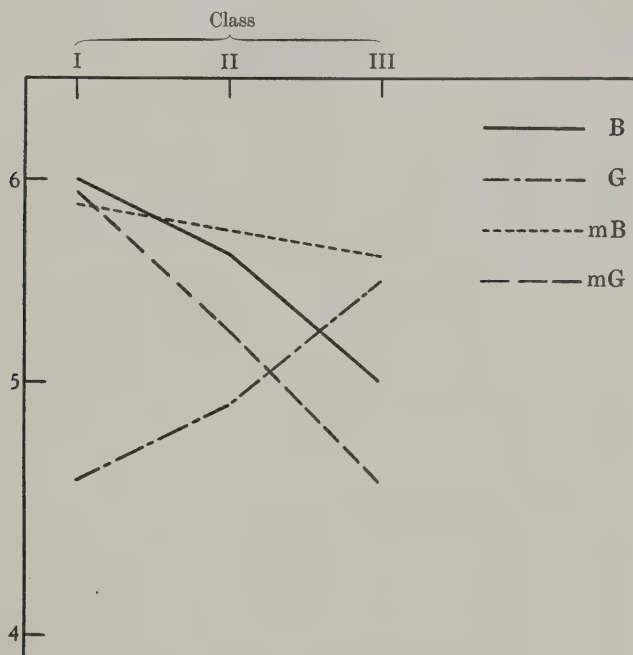


Fig. 1. Test of understanding. Means.

III. TEST RESULTS

It was originally hoped that scores on the tests used would vary directly with age, and that they might be at least comparatively independent of schooling. Nevertheless, the results were more easily classifiable, in the first instance, according to classes, and they were therefore treated accordingly. Means were calculated for each class separately, and these are shown in graphical form in Fig. 1 (Understanding) and Fig. 2 (Taste).

(1) *Test of understanding*

The results for the test of understanding are also shown in Table IV, together with mean scores on the same ten test items for two adult groups studied in the previous experiment. It will be seen that in three

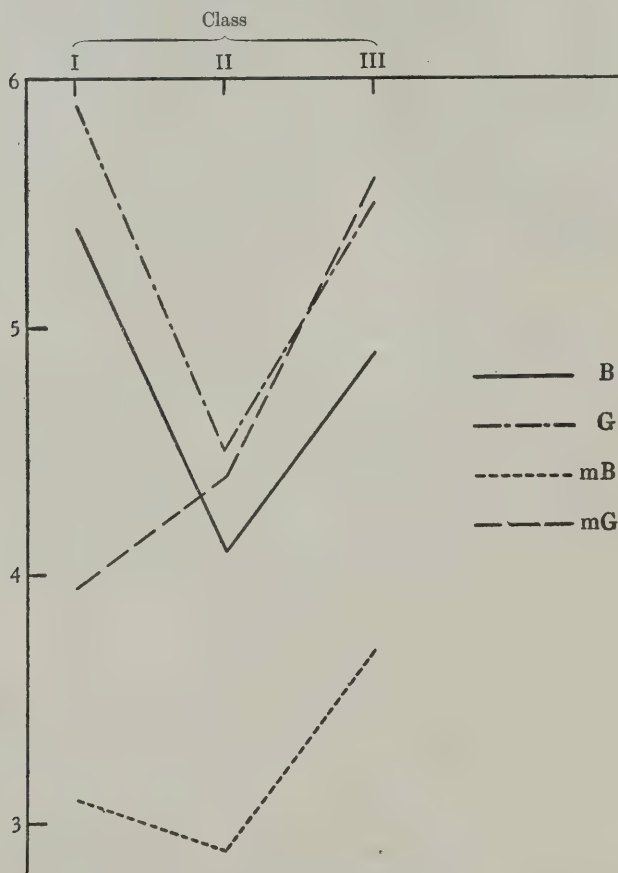


Fig. 2. Test of taste. Means.

groups out of four there is a continuous and fairly regular drop in scores as one progresses from class I to class III. In the fourth group (the girls' school) the curve takes the opposite course. Redistributed into age groups the scores do not vary so regularly. This redistribution was carried out in two different ways: (a) according to *age last birthday*, and (b) to *age correct to one year*. The resulting curves are not included in the

Table IV. *Test of understanding: comparative scores of juvenile and adult groups*

Group	Class			All
	I	II	III	
B	5.8	5.5	5.0	5.6
G	4.6	4.9	5.4	4.9
mB	5.7	5.6	5.5	5.6
mG	5.75	5.2	4.6	5.0
Unselected adults	—	—	—	5.6
Music students	—	—	—	7.1

present report, but it may be remarked that the curves for groups 'B', 'mB', and 'mG' still showed a decided tendency to fall with rising age, while that for group 'G' plotted according to method (a) rose much less steeply, and plotted according to method (b) showed an irregular up-and-down movement. The net result would seem to be that success in this test is less a function of age than of schooling. Nevertheless, correlations between test scores and age, shown in Table V, are interesting, and may be summed up by saying that on the whole there is a slight but definite tendency for musical understanding, as measured by this test, to deteriorate with age, although a certain degree of recovery occurs as adulthood is reached.¹

Table V. *Test of understanding: correlations of scores with age*

Group	$r \pm \text{P.E.}$	Remarks
B	-0.28 ± 0.066	Significant
G	$+0.19 \pm 0.074$	Not significant
mB	$+0.04 \pm 0.093$	Approx. zero
mG	-0.42 ± 0.081	Significant
All	-0.17 ± 0.040	Significant

(2) *Test of taste*

Curves for the test of taste corresponding to those for that of understanding are shown in Fig. 2. As in the previous case, three of the four curves show the same general outline—in this instance *down*, between classes I and II, and then *up*, between classes II and III. The fourth curve also shows a rise in this second portion, but this is preceded by another less steep rise. Redistributing into age groups, as before, one finds the same general form preserved, more or less, for the three groups showing uniformity, and also, when method (a) was used, in the fourth group (mG). In view of these facts it is not surprising that the overall

¹ It should further be noted that partialling out 'knowledge' increases the negative correlation between understanding and age (see Table VI).

154 *A new approach to the testing of musical ability*

correlation between age and taste scores is practically zero (actually -0.017 ± 0.042). Taking the groups separately, none of the correlations is significant one way or the other. The important point, however, is that while from the age of 14 or thereabouts a gradual approximation to adult standards of taste takes place, a very definite fall occurs before that age is reached.

(3) *Test of musical knowledge*

It is not proposed to discuss in detail the results of the test of musical knowledge. Reference to Table VI will show a fairly low, but highly significant correlation with age. The scores obtained ranged from 12 to 33, with an overall average of 22.6. A fairly steady rise was found with progression through both classes and age groups, except that the youngest children tended to score slightly higher than their immediate seniors. This, however, is easily explicable in terms of the conditions of sampling; without this complication the correlation noted above might have been higher.

Table VI. *Correlations between all measures, juvenile and adult groups*

Measures correlated	$r \pm \text{P.E.}$	Partial r	Adults $r \pm \text{P.E.}$
AU	$-0.17 \pm 0.040^*$	$r_{\text{AU.K}} - 0.24^*$	—
AL	-0.12 ± 0.042	—	—
AT	-0.017 ± 0.042	—	—
AK	$0.31 \pm 0.041^*$	—	—
UL	0.027 ± 0.041	—	0.15 ± 0.056
UT	-0.041 ± 0.042	—	—
UK	$0.16 \pm 0.045^*$	$r_{\text{UK.A}} 0.22^*$	$0.39 \pm 0.052^*$
LT	$0.17 \pm 0.042^*$	—	—
LK	0.001 ± 0.046	—	0.02 ± 0.059
TK	$0.22 \pm 0.044^*$	$r_{\text{TK.A}} 0.24^*$	$0.73 \pm 0.080^*$

A, age; U, understanding; L, 'liking'; T, taste; K, knowledge.

* Significant coefficients (coefficients not less than three times their probable errors).

(4) *Correlations between tests*

Table VI summarizes the correlation coefficients already mentioned, together with others it was thought desirable to compute. Reference to the writer's previous paper (4) will show that in the case of that experiment it seemed desirable to take into account a further measurable variable. This consisted of the number of selections in the test of understanding which the listener stated that he liked (irrespective of qualifications as to the degree of liking). The term 'liking' was used to express this number, and it was suggested that this represented a measurable quantity which was relatively constant for the individual. 'Liking' scores were therefore

counted in the present instance as well, and correlations of these scores with the others are included in Table VI, together with any available comparable figures from the experiments with adult subjects.

The following main results, in addition to those already noted, emerge:

(1) Both understanding and taste show a positive correlation with knowledge, which, though low in each case, is nevertheless significant. Partialling out age, incidentally, increases the correlations to a slight extent.

(2) Correlations between understanding and taste and between liking and understanding are both practically zero. That between liking and taste, on the other hand, is positive, and of the same order as those discussed in the previous paragraph.

(3) The net result seems to be that understanding and taste are independent musical abilities. Whether 'liking' measures another aspect of musical talent is more open to doubt, since, unlike the other two measures, it has a zero correlation with knowledge, the only external criterion available within the limitations of the experiment.* (The validity of knowledge scores as a criterion of musical ability is briefly discussed later.) If, however, its status could be established, it would appear that 'liking' would also have to be regarded as an aspect of musical ability more or less independent of the other two.

(5) *Sex differences*

A brief summary of sex differences observed in the test results is given in Table VII.

Table VII. *Sex differences in test results*

Test	Boys		Girls		Diff.	P.E. diff.
	Mean	No.	Mean	No.		
Understanding	5.51	141	5.01	123	0.50	± 0.100
Taste	4.31	136	5.06	122	0.75	± 0.192
Knowledge	23.2	98	22.1	89	1.1	± 0.498

It will be seen that the boys show just about the same relative superiority over the girls in the test of understanding as the girls do over the boys in that of taste. This slight inferiority of the boys on the more purely aesthetic side confirms popular opinion regarding sex differences, as does their slight superiority in the informative knowledge test. (In calculating means for this latter the class I girls corresponding to the class I boys who did not take this test are omitted.) The differences in

respect of understanding and taste are fully significant; that for knowledge, however, is only of the magnitude classed as 'suggestive'.

IV. DISCUSSION

(1) *Test of understanding: responses of children of different ages*

In the light of the high score obtained by young children in the understanding test, it is obvious that the applicability of the method is not confined to adults. How far down the age scale it would be possible to go, further adapting the material, is uncertain, but it is interesting to note that in the case of two items, no. 7 and, more particularly, no. 9, the percentage of correct answers recorded was higher in the juvenile than in the adult group. In only two instances, nos. 4 and 5, was the proportion of successes appreciably higher in the adult group. These, incidentally, are the two items from the *Carnival of Animals*, commonly described as specially suitable for children. In view of these findings it is of some interest to try to discover whether success in the individual items tended to increase with age. To this end, analyses into year groups were made in the same way as for the total test scores, with the results shown in Table VIII.

Table VIII. *Test of understanding: variation in understanding of individual items with age*

	Direction of curve with increasing age				
	Up	Down			
Boys	6	2	3	5	10
Girls	—	5	8	10	

It will be noticed that the downward tendency is much more common than the upward, and that it occurs not only in respect of rhythmic items (e.g. no. 5), but also in respect of music of a definitely adult character (no. 10).

Responses regarding 'liking' were also analysed, but without any result of outstanding interest. There were few marked differences between the respective proportions of children and adults who liked any individual item. The most marked divergences were in respect of no. 4, in which the percentage of adults was higher, and of no. 9, where the percentage was lower than that of either boys or girls in any single year group. Fairly steady drops in popularity were observed, for the boys in respect of

nos. 1, 2, 4 and 5, and for the girls in respect of nos. 1 and 10. Regular variation in the contrary direction did not occur at all.

The net conclusion would seem to be that within the limits of the age groups studied, the material used showed no special unsuitability for children of any age, except the rather curious fact already hinted at, that music commonly considered specially suitable for children was neither more correctly interpreted nor more popular than the other selections used. It would seem to follow that at least in small doses children can appreciate much the same type of music as adults, and to the same degree.

(2) *Test of taste: validity of instructions and of
general method*

Much more marked differences were found in the scores obtained by the juvenile and adult groups in the test of taste. In drawing up the test the writer had experienced some misgivings regarding the instructions, especially in relation to adult subjects. In his original application of the technique Mohler had used the terms 'best', 'next best', and 'poorest'. Since it was felt that at least with adult subjects the music judged 'best' might not necessarily be the same as that most enjoyed, the instructions were altered to the form in which they are quoted on p. 147 above. The writer had anticipated objection to the method on this score, but since no such objections were put forward, no questions as to whether such a thought had occurred were asked, in case an opportunity of further investigation of this point might arise. However, since the placings of the records by majority vote agreed with the assumptions regarding merit on the basis of which the selection was made, it seemed that the potential difficulty had not arisen—provided that strict adherence to instructions had been maintained on the part of the subjects, an assumption that it was fairly safe to make.

The serial order of the groups in the test was intended to present progressively finer degrees of difference in relative merit. It may be of interest to note that the results were contrary to what might have been expected: higher scores were obtained by the children on group III than on group II, and on group II than on group I. This may have been caused in part by the high rhythmic appeal of records IB and IIc, which were graded (3) in each case by the adults. This raises the point of the degree to which works in each group could be compared on terms as it were of equality. Since the records were chosen from the rather limited number which happened to be readily available, the selection was felt to be far from satisfactory. Nevertheless, it was found possible to include in each

group works of approximately the same tempo and general emotional character.

(3) *Suitability of technique for the groups tested*

As regards the suitability of the actual testing technique, it is felt that a test period of 40 min. was not too long for children of the ages studied. On the other hand, the average playing time of a 12 in. or even a 10 in. record seems to be too long for sustained attention, especially in the case of the taste test, where a rather difficult attitude of combined expectancy and recall had to be maintained. Where possible, less than the whole of a single side was played, but the adoption of this expedient is undesirable in other ways, e.g. it may destroy the balance of a work such as a short classical overture. Shorter selections would allow of a larger number of groups for comparison in the same time, and this would make for both increased interest and increased reliability.

V. GENERAL CONCLUSIONS

In addition to the empirical results already discussed, the present experiment suggests the following general conclusions:

(1) The investigation yields no very definite evidence one way or the other regarding the possibility of establishing tests of untaught musical ability. The uniformity of the 'class' curves seems to indicate that a very considerable effect is produced by differences of schooling. On the other hand, the fact that variations with age did emerge, and that these included a marked falling-off in test scores in certain portions of the age range studied, suggests that under other circumstances one might be able to eliminate the effect of education to a greater extent. (This is borne out by the effects of partialling out certain variables, as already noted.) 'Other circumstances' which would repay study include groups of still younger children, and of persons who have had no training whatever in music or musical appreciation.

(2) The statistical treatment of the present data suggests that it is impossible to postulate the existence of unitary musical ability, except in so far as the talented musician would be regarded as one who possessed in a high degree all or most of a constellation of abilities, each of which had its significance in the performance, interpretation, or enjoyment of music. Accordingly, in music testing a 'profile' technique would seem to be indicated in preference to a pooling of results into a total score. A method such as this has been suggested by Schoen⁽²⁾, who used his own battery of tests⁽³⁾ as a basis for prognosis of the type of musical activity

for which a given individual was fitted. Such a 'profile' might also include results of tests of an active and for some purposes of a creative character (e.g. those of Vidor⁽⁶⁾), which fell outside the scope of the present investigation. It is hoped later to add to the present battery a test of the appreciation of form.

(3) The main difficulty in drawing up any such battery is that of determining the validity of any test as an indicator of musical ability or an aspect thereof. In the present instance, the only possible criteria were 'knowledge' scores and teachers' estimates. In his previous paper the writer argued that other things being equal persons musically gifted would show more interest in matters musical, and consequently achieve higher scores, but this is admitted to be at best a slender argument, especially when applied to the juvenile groups. Teachers maintain, with some degree of justice, that they are able, through personal contacts, to point to the gifted individuals in a particular group, but something more than this is necessary. An attempt was made by the writer to discover whether mean scores on all measures obtained by children who claimed to play an instrument were significantly higher than scores obtained by children who did not; the schools were taken both as a single unit and, since there was considerable inequality of opportunity, separately. No significant differences were found for separate schools, but taking all together every measure with the exception of understanding gave means for 'players' significantly higher than for 'non-players'. These differences were as follows: liking, 0.67 ± 0.166 ; taste, 0.70 ± 0.196 ; knowledge, 1.84 ± 0.387 . (Understanding gave a margin of 0.17 in favour of the non-players.) A variety of difficulties, such as knowing how to deal with performers on the Jew's harp and Kazoo, and the heterogeneity of the sample already noted, make this criterion one of doubtful value in the present case, but the observed differences are highly suggestive. A combination of all three criteria might be used to give a sufficient indication of validity.

(4) While tests of the additional types mentioned in paragraph (2) above are of undoubted importance, it is felt that since at present the musical experience of the majority of people is limited to listening to music, attention should be paid to methods of testing which give the closest possible approximation to the normal experience of listening to music. Normal modes of listening vary considerably with circumstances, e.g. listening to a 'live' performance in a concert hall is a very different experience from listening to the gramophone or radio by oneself or in the company of a few friends, and the test situation in the present experiment

had its differences from and affinities with both. While the practice of having a large group of subjects listening simultaneously has its obvious advantages, the larger the group the more distractions, as in all group testing. Since, however, the method described takes up a good deal of time whatever the circumstances, it is felt that the risk of error may be accepted, and a group of moderate size be tested as a unit. Useful results from testing this variety of musical experience are possible not only from the point of view of the person tested, but also with regard to the light that they throw on the scope of the average listener's potential interpretative ability and aesthetic discrimination. It is hard to believe that the mediocrity of the musical entertainment commonly presented represents the limits of what the ordinary man or woman can be expected to appreciate.

(5) It is not suggested that fine differences of any aspect of musical talent are measurable by the method employed. Other investigators who may be interested to repeat the experiment might profitably make their own selection of records, and standardize within whatever degree of accuracy they think fit. In particular, new and improved series of test records could be made by a more complete exclusion of relatively familiar music, the use of which has its obvious disadvantages. Accordingly, the only claim put forward for the tests described is that they afford a means of detecting extremes of endowment in respect of certain qualities of a musical character. Whether there exists in addition a quality of 'musicality', testable by other means, remains to be seen.

VI. SUMMARY

The experiment described continues a line of enquiry reported in a previous paper, by extending a test of musical understanding there described to children between the ages of 12 and 15 inclusive, and by adding a test of musical taste, the standards for which were obtained from the majority vote of a small group of adults interested in music. In addition, a new test of musical knowledge was applied to the same children.

It was found that in general test scores did not increase with advance in age, but that, particularly in the test of understanding, the opposite took place. Scores on the two main tests were found to be uncorrelated, but each test correlated positively with knowledge. Significant sex differences were observed—boys having the advantage in the test of understanding, girls in that of taste. Significant differences on most

measures were also found between children who played an instrument and those who did not.

The method and material both appeared to be as suitable for the juvenile groups studied as for adults. No conclusive evidence, however, was obtained regarding the extent to which the scores were affected by differences in education. Since there seems so far to be no reason to believe in the existence of a unitary musical ability, it is suggested that results of separate tests should not be pooled but presented in the form of a profile, possibly with others of a more active character. It is further suggested that while the technique employed is both valid and useful, other investigators who may care to repeat the experiment should be able to frame test series consisting of other records equally or more suitable for the purpose.

I should like to express my thanks to the headmasters and head-mistress of the schools concerned for permission to carry out the testing programme, and my appreciation of the friendly co-operation received from the music teachers, from the children themselves, and in particular from the students and others who acted as voluntary subjects in the check experiment described in section II (2).

REFERENCES

- (1) KWALWASSER, J. (1927). *Kwalwasser Test of Music Information and Appreciation*. State University of Iowa, Extension Division.
- (2) SCHOEN, M. (1923). "The validity of tests of musical talent." *J. comp. Psychol.* III, 101-21.
- (3) — (1925). "Tests of musical feeling and understanding." *J. comp. Psychol.* v, 31-52.
- (4) SEMEONOFF, B. (1940). "A new approach to the testing of musical ability." *Brit. J. Psychol.* xxx, 326-40.
- (5) TRABUE, M. R. (1923). "Scales for measuring judgement of orchestral music." *J. educ. Psychol.* xiv, 545-61.
- (6) VIDOR, M. "Examining musical ability." (Paper read at Extended General Meeting of the British Psychological Society, Reading, April 1939.)

(Manuscript received 14 December 1939)

DIP IN WORK CURVES

By D. F. WARD

From the Psychological Laboratory, University College, London

- I. *Introduction* (pp. 162-163).
- II. *Procedure* (pp. 163-164).
- III. *Results* (pp. 164-170).
 - (a) *Appearance of work curves* (pp. 164-166).
 - (b) *Measurement of dip* (pp. 166-167).
 - (c) *To find a measure of increase in amplitude* (pp. 167-169).
 - (d) *Correlation between dip and increase in amplitude* (p. 169).
 - (e) *Amplitude and trough sequences in the dip region* (pp. 169-170).
- IV. *Summary* (pp. 170-171).
 - References* (p. 171).

I. INTRODUCTION

THE results here presented were obtained during a study of fatigue effects and their use as measures. The conventional measure of fatigue is that of loss in output over a given time, sometimes called the *decrement*. Another common suggestion is that tired work is more variable or erratic, whence the notion of a measure in terms of *increase in variability*, i.e. in the amplitude of the waves in the curve. The first problem of this present research was to see if there is a correlation between different forms of these measures⁽¹⁾.

When plotted on ordinary squared paper, a work curve is roughly hyperbolic. It has, therefore, been suggested that it should plot out approximately as a straight line on two-way log paper⁽²⁾. If this is so, the simplest measure of decrement would be the downward angle or slope of that straight line. This is what we shall call the *angle of decrement*.

Matters were complicated during this research by the fact that the curve was still slightly curved, even when plotted on two-way log paper. A ruler laid on a typical curve, to join its first¹ and last points, may have all the rest of the curve below it in a slight 'bow'. The point of maximum departure lies somewhere between the 300th and 600th sec. in curves of

¹ As will be noted in the text, the actual first point of a curve was always ignored, so that the effective first point here mentioned was really the second. It is a small matter, but is mentioned for the sake of complete accuracy.

a total length of 1200 sec. Here, the curve may in fact cease to fall, curving less and less until it is horizontal, or even rising slightly. We thus arrived at the need for a third measure, that of *dip or downward bulging* as distinct from those of *decrement* over the whole curve and of *increasing variability* respectively.

II. PROCEDURE

Eighty students from training colleges acted as unpractised subjects for a Kraepelin adding test, adding two to each number along a row, working as quickly as possible, and beginning a fresh row every 5 sec. The output or work done by each subject was measured by the number of figures correctly computed each 5 sec. The subjects were composed of three groups, working for 1200, 1500 and 1800 sec. respectively, the outputs per 5 sec. for each subject were pooled for 25 sec., and the results plotted on two-way log paper, time being represented horizontally and output vertically.

It has been previously shown by Philpott⁽²⁾ that "the curve of work has a general form roughly approximating to a rectangular hyperbola, and that, if the logs of time and output are plotted, the work curve should tend to appear as a straight line, the fatigue decrement thus being measurable in terms of a single linear expression".

The rectangular hyperbola can be expressed by the equation,

$$xy = k \text{ (a constant),}^1$$

and taking logarithms,

$$\log x + \log y = k' \text{ (a constant).}$$

Hence if $\log y$ be plotted against $\log x$ a straight line will result and it will have a negative gradient of unity.

If the curve is a perfect rectangular hyperbola, by plotting log output against log time it will be translated into the straight line form. Actually, of course, this is not so, but if the curve of output does fairly approximate in shape to such a hyperbola, the plotting of the logs of time and output should effect a considerable flattening.

We found that work curves of individual subjects so plotted were considerably flattened and smoothed, and that the grand total curves were approximately straight lines within certain limitations.

In finding a measure of decrement the first point on each curve was omitted. The reason for this is as follows:

Let it be supposed that output is constant in log time, i.e. that the subject does ' p ' units of work in every successive time interval. The

¹ So far as is known to the writer, Birtwistle⁽³⁾ was the first to use this new technique.

resulting curve of output will be a horizontal straight line in log time. If the line is now expressed as output per second in ordinary time, it will take on a hyperbolic appearance, although it will not be a true hyperbola as it stands, being amounts per finite instead of infinitesimal time units, and further it will not be asymptotic about the axes of time and output.

Now let this second variant be re-expressed as a curve on two-way log paper. It will be almost a straight line, save for the first point, and to a very slight extent, the second point. But it will no longer be horizontal; it will be inclined to the base line, as has been previously shown.

Our measure of decrement was the line of 'best fit' to the remaining portion of the work of each subject. This was calculated by finding a regression line. Using Yule's notation for regression lines⁽⁶⁾, the line suitable for the purpose is

$$y = b_2x, \text{ where } b_2 = r \cdot \frac{\sigma_y}{\sigma_x},$$

and for a work curve x will be a time measure and y an output ordinate. Further, b_2 is the tangent of the angle of inclination of the line of 'best fit' and could serve as a satisfactory measure of decrement.

In practice, twenty-one equidistant points are marked on the time line, each 1 cm. apart. From each point an ordinate, measured in millimetres, is drawn to the curve; the mean of these ordinates is calculated. The x values are calculated as arbitrary units, -10 to $+10$, and the y values are deviations with respect to sign from their mean (actually in millimetres). Thus the regression line is of the form

$$y = \frac{\Sigma xy}{10\Sigma x^2} \cdot x \text{ and } b_2 = \frac{\Sigma xy}{10\Sigma x^2}$$

and will have a negative value.

A separate work curve was plotted for each subject, and a regression line evaluated from each curve.

III. RESULTS

(a) *Appearance of work curves*

Studying the individual work curves, in each case there appeared a dip or concavity (see Fig. 1). The region of the dip in every curve was from 400 to 800 sec. approximately, this dip being more marked in some cases than others. After 900 sec. the curves rose steadily in nearly every case.

It is seen in Fig. 2 that our grand total curves fitted the hyperbolic theory extraordinarily well for the first 400 sec. From this point there

was an abrupt change. The slope vanished and the curve was practically horizontal, and from 600 sec. the curve gradually ascended for the remainder of the time. The part after 600 sec. appeared much more erratic than the downward slope for the first 400 sec., owing partly to the crowding of points.

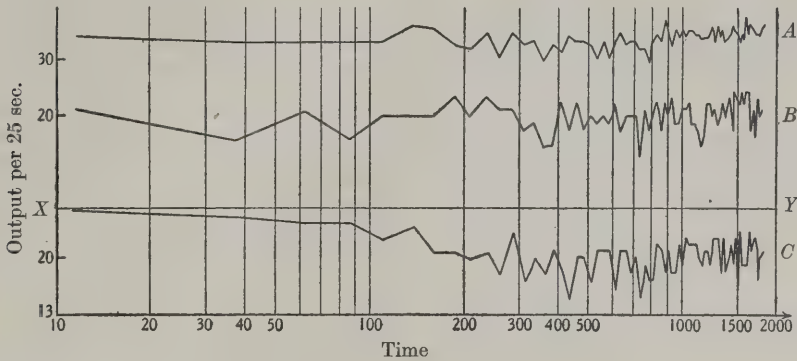


Fig. 1.

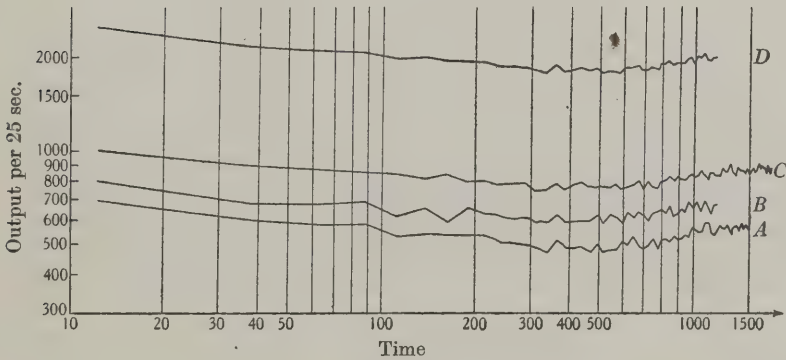


Fig. 2.

Thus our line of 'best fit' only fits well for the first 400 sec. or the first 600 sec. at most. The decrement thereafter tends to be lessened by the rise in the curve, the longer the curve the greater the lessening. In fact, the output is almost as great at the end of the curve as it is at the beginning. If we measured fatigue by the old method of determining the difference between initial output and final output, very little fatigue would be shown.

The effect of the dip and the following rise in the curve can be made manifest by inquiring what kind of measures we should have obtained had the subjects stopped work earlier. Our experiments actually lasted 1200 sec. What measures of decrements, etc., would there have been had they lasted only for say 100 or 200 sec., etc.?

For purposes of this inquiry, the curves were plotted against the log of time. The base line was then arbitrarily cut into twenty-one equal spaces (logarithmically speaking). Regression lines were then calculated over the first nine intervals, over the first eleven, the first thirteen, and so on. We thus obtained the measures of decrement that would presumably have been obtained had the original experiment been cut to one or other of the given lengths.

The figures are given in Table I against the (arbitrary) log lengths concerned together with the corresponding lengths in seconds.

Table I

Arbitrary log measure	Corresponding number of sec.	Tangent of angle of decrement
9	195	-0.0650
11	285	-0.0727
13	410	-0.0780
15	590	-0.0711
17	855	-0.0566
± 19	1200	-0.0388

From this table we observe that decrement is fairly constant for as far as 600 sec., i.e. to the extent of the dip. After this, the rise in the curve affects this value by compensating for the dip, i.e. lessening the value of the decrement. It is noticed that the value of decrement for the first 400 sec. is twice that of the 1200 sec. If we calculated for 1800 sec. the value would probably be still lower owing to the upward trend of the curve.

In stating our problem we stated that decrement was one of our measures of fatigue. Hence, from the above table, we see that to some extent it is unreliable. The measure is fairly constant so long as the curve has a downward slope, i.e. for 600 sec. Thus the actual value which we attach to decrement depends upon the length of the curve.

(b) *Measurement of dip*

At this point we can attempt a measure of dip. There are two main angles, that over the whole curve already measured (covering say 1200 sec.), and the angle over the first portion of curve, say up to 600 sec.

In nearly every case, the slope of the 600 sec. line was greater than that of the other, and by an amount that can be given as $b_2 - c_2$, where b_2 and c_2 are the tangents of the two angles concerned; or of course, it can be expressed in degrees and minutes, this giving the angle of dip. The angle of dip is then a difference. It is the amount by which the slope varies when we calculate it over 600 and 1200 sec. Other measures would have been possible. This one will suffice. It at any rate gives some measure of the thing desired to be measured, namely departure from simple linearity. It is conceivable that this departure would occur more readily in cases of extreme fatigue. That is to say, there may perhaps be a correlation between dip and the decrement (i.e. the 1200 sec. angle). The obtained value was +0.26. There is then some degree of relation.

By the method of calculating the regression line for decrement, more weight is given to the extreme points, i.e. beginning and end of the curve. Thus large values at the end of the curve could make our value decrement seem small. Conversely, when calculating the regression line for the first 600 sec. only, more weight is given to the end of that particular curve, i.e. to the low value of the dip region. This will give the regression line for the first 600 sec. a large decrement, and possibly a correspondingly large value for the angle of dip. This accounts for the only moderate value for the correlation coefficient.

(c) *To find a measure of increase in amplitude*

Several difficulties had to be overcome in obtaining this measure. We here give an outline of them.

Amplitude represents the difference in output from a crest to the next succeeding trough—or *vice versa*—in a work curve. As there are so many differences in a curve of 1200 sec. we must find a method which takes into account all of them, at the same time not giving undue weight to a particular section of the curve. Flugel's curve of successive differences⁽⁴⁾ does show this, but we set out to find another measure.

Our first attempt was found to be untrustworthy. If we take the measure of decrement for the 1200 sec. utilized above, the regression line is really a line of ideal output throughout the length of the test. This we can utilize, plotting each point of a work curve as a deviation, regardless of sign, positive or negative, from the amount of ideal output at that point.

Taking the deviations for the vertical axis, and the time as the horizontal axis, we can plot on ordinary graph paper a graph for each subject.

Similarly, utilizing the method outlined for calculating a regression line, we can obtain a line of 'best fit' for the subject's deviations. The slope of this line, with regard to sign, could be taken as a measure of increase in amplitude.

Our line of 'best fit' which we have taken for decrement although affected by the dip factor and the upward trend from 900 sec. onwards, cannot wholly take these factors into account. Thus our deviations in the centre of the 400-800 region will have too much weight, whilst at the 800 region they will have too little weight, and for the regions afterwards they will tend to have undue weight. From these findings it will be seen that the deviations will tend to give an undue or incorrect increase in amplitude, especially as the rise after 900 sec. gives added weight to the regression line.

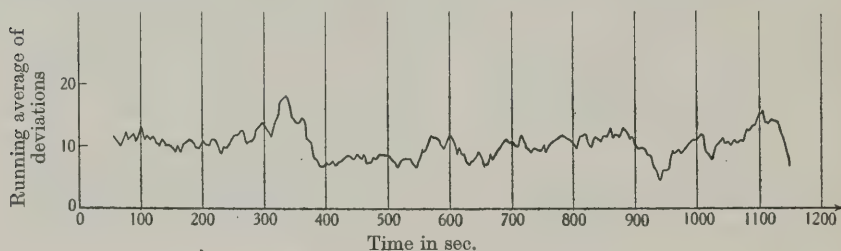


Fig. 3.

In our second attempt we utilized a curve of running average. "To get the curve of running average, take the average of points 1 to 11 in the original curve, then the averages of points 2-12, 3-13, and so on. These successive averages plotted against the centres of the intervals concerned give the desired curve" (5), p. 112). Every point in the original curve was now expressed as a deviation from the running average, plotting without regard to sign. From this graph of deviations, the regression line was found in the usual way. The slope of the line was our measure of increase in amplitude.

As a curve of 1200 sec. would have nearly 240 points, we marked each fifth point on the time or x axis, and used an arbitrary time scale from -23 to $+23$ (i.e. 47 points). We measured the y values for these particular points. Actually in plotting these curves we plotted the deviation between the total running averages (i.e. not divided by eleven) and eleven times the output for the said point. We further produced a grand total curve of deviations from running average. We smoothed this curve by taking the running average of it (see Fig. 3). We found there is a rise in the curve

as far as 330 sec., but this is followed by a very sudden fall, corresponding to the beginning of the dip region. During the dip region, from 400 to 550 sec., the curve is more or less horizontal. On leaving the dip region, i.e. from 550 sec. onwards, the curve shows an upward trend although this is not constant, taking place by a series of rises and falls. It is noticed that this rise is gradual compared to the abrupt fall from 330 to 400 sec.

In our investigations, we considered Flugel's measure of increase in amplitude, i.e. successive differences. His measure appears to be wholly satisfactory for his experiment. His subjects became much more practised as the experiment proceeded daily, and hence no dip region was noticed in his work curve. Utilizing his measure for our own experiment would not eliminate the dip factor, in fact it would accentuate it, making the correlation untrustworthy in the same way that we found to occur with our first measure, i.e. measuring deviations from the line of 'best fit'.

(d) Correlation between dip and increase in amplitude

In the research we correlated increase in amplitude with the measure for decrement for the whole 1200 sec.; our result was -0.14 . To show the effect of dip in our curves we correlated the measure for decrement for the first 600 sec. with the measure for increase in amplitude for the whole 1200 sec., the correlation being -0.075 , showing there is little relation between these two measures. Further, we set out to ascertain whether there was any correlation between angle of dip and increase in amplitude. We again only obtained a low value of -0.042 .

(e) Amplitude and trough sequences in the dip region

Concerning the dip region we find two interesting points. The deviations are only more noticeable at the beginning and end of the region. During the period of low output the deviations are almost constant, and in fact they are practically the lowest for the whole curve.

Thus it seems that increase in amplitude is most marked when the depression or conflict, as mentioned in the subjects' introspections, is taking place or asserting itself. When the subject's depression has manifested itself, his output is low, but the amplitude is correspondingly low. When his output increases, i.e. after about 600 sec., amplitude increases with it.

Philpott⁽²⁾ has shown that work curves have zones *A* and *B*, and that the wave-length in zone *A* may differ from that in zone *B*, and also that

the L.C.M. sequences are of a composite nature. The curves have many L.C.M. cycles, so many in fact that those characteristic of zone *A* can summate to accentuate the troughs of that zone, cycles of zone *B* pattern summing in their turn, the resulting curve thus seeming to change in wave-length as we pass from one end to the other. This at least usually occurs, though not always.

For a grand total curve he has shown that troughs should occur approximately at 49, $63\frac{1}{2}$, 80, 103, 130 ... sec. Considering our own grand total curve we find that troughs occur at $63\frac{1}{2}$, 80, 103, 130 ... sec.

Studying work curves we find that amplitudes are small to begin with, afterwards opening out, but then there comes a period of small oscillations followed by the amplitudes sweeping out. This period of small oscillations corresponds to our dip region. Philpott has shown that this period of small oscillations is the period of change from zone *A* to zone *B*. This change is quite noticeable in our grand total curve. It may perhaps be considered due to the beat effect of L.C.M. cycles.

IV. SUMMARY

Dip or downward bulging has occurred in all our work curves, being confined to a particular region of the curves, from 350 to 600 sec. We have measured it by determining the difference in the angle of slope of the two regression lines: those taken between 0 and 1200 sec., and from 0 to 600 sec. respectively.

There is a small positive correlation between the angle of dip and the angle of decrement, showing that if angle of decrement was large, output in the first 600 sec. might have decreased very rapidly.

On the other hand, dip seems to have little connexion with increase in amplitude over the full 1200 sec., nor does the regression line for the first 600 sec. correlate to any extent with increase in amplitude. So that rapid drop in initial output does not necessarily show that output in the latter part of the test will be erratic.

Amplitudes of the waves of the curve were found to be small in the dip region (from 350 to 600 sec.), but were large just before and after it. Again, dip takes place in the region where zone *A* is merging into zone *B*; further, because the dip occurs in this region, amplitudes are small.

Also the beginning of the dip region is marked by the manifesting of feelings of conflict; on this dying down, the rate of output increases, and increase in amplitude is also observed.

It may be that the conflict is due to change in the zones and to the

smallness of the amplitudes, or, the other way round, the switch from zone *A* to zone *B* may be due to the conflict. This problem has not been solved by the enquiry.

REFERENCES

- (1) WARD, D. F. (1939). "Correlation between increase in amplitude and decrement in work curves." Unpublished Thesis, University of London Library.
- (2) PHILPOTT, S. J. F. (1934). "Theoretical curve of fluctuations of attention." *Brit. J. Psychol.* xxv, 221.
- (3) BIRTWISTLE, J. A. (1938). "The work curve and the work decrement." *Brit. J. Psychol.* xxviii, 315.
- (4) FLUGEL, J. C. (1928). "Practice, fatigue, and oscillation." *Brit. J. Psychol. Monog. Suppl.* no. 13.
- (5) PHILPOTT, S. J. F. (1932). "Fluctuations in human output." *Brit. J. Psychol. Monog. Suppl.* no. 17.
- (6) YULE, U. (1932). *An Introduction to the Theory of Statistics*. London: Griffin.

(*Manuscript received 30 January 1940*)

A BI-FACTOR ANALYSIS OF RELIABILITY COEFFICIENTS

BY GEORGE A. FERGUSON

*From the Education Department, Moray House,
University of Edinburgh*

- I. *Holzinger's bi-factor method* (pp. 172-173).
- II. *A bi-factor analysis of the intercorrelations between the halves of three equivalent test forms* (pp. 173-175).
- III. *Interpretation of factors* (pp. 176-177).
- IV. *A comparison of multiple orthogonal factors with bi-factors* (pp. 177-180).
- V. *Some observations regarding the comparison in section IV* (pp. 180-181).
- VI. *Summary* (pp. 181-182).
- References* (p. 182).

I. HOLZINGER'S BI-FACTOR METHOD

HOLZINGER'S method of bi-factor analysis⁽¹⁾ attempts to describe a matrix of correlations in terms of one general factor, a number of group factors common to two or more variables, and as many specific factors as there are variables. This reduces the matrix to a minimum factorial description of one general factor, n specific factors, where n is the number of tests, and q group factors, q being smaller than n . The procedure is to examine the matrix of correlations to be factorized in order to isolate any groups of tests that correlate more highly among themselves than they do with the remaining tests in the battery, grouping those tests together whose intercorrelations constitute elements in vanishing tetrads.

In allocating tests to groups Holzinger uses what is termed a *B*-coefficient. A *B*-coefficient is defined as, "the average of all intercorrelations of tests 1, 2, ..., K , divided by the average of all correlations of tests 1, 2, ..., K , with the remaining tests not in the group"⁽²⁾.

Having allocated the tests to groups, the next procedure is to remove the general factor. This is accomplished in a manner similar to that employed by Spearman in estimating g loadings by averaging all possible combinations of

$$r_{ig}^2 = \frac{r_{ij}r_{ik}}{r_{jk}}. \quad (1)$$

In the bi-factor method, however, only those values of r are used that are elements in tetrads approximating zero.

Let the following represent a hypothetical bi-factor pattern with six variables:

	A	B	C	D
1	a_1	b_1	.	.
2	a_2	b_2	.	.
3	a_3	.	c_3	.
4	a_4	.	c_4	.
5	a_5	.	.	d_5
6	a_6	.	.	d_6

Examination of this factor pattern will show that certain tetrads such as $r_{13}r_{24} - r_{14}r_{23}$ will be zero, while certain others such as $r_{12}r_{34} - r_{14}r_{23}$ will be greater than zero. In the above factor pattern there will be four values of r_{ia} , which with fallible data must be averaged. Thus the formula for the general factor loading of the first variable becomes,

$$r_{1a}^2 = \frac{r_{13}r_{15} + r_{14}r_{16} + r_{13}r_{16} + r_{14}r_{15}}{r_{35} + r_{46} + r_{36} + r_{45}}, \quad (2)$$

The general factor having been removed, a table of residual correlations is calculated, and the group factors removed successively.

II. A BI-FACTOR ANALYSIS OF THE INTERCORRELATIONS BETWEEN THE HALVES OF THREE EQUIVALENT TEST FORMS

In the present enquiry three Moray House Group Tests of Intelligence M.H.T. 21, M.H.T. 23, and M.H.T. 26, were given to a sample of 212 children of 11+ in West Yorkshire schools.

Each test contained 100 items and required 45 min. to administer. One week elapsed between the administration of each test. The three tests are similar in structure and are regarded as parallel forms. The scores of each child on the odd and even items of each test were found. The fifteen different intercorrelations between the six halves of the three tests were calculated. Three of these intercorrelations are between halves of tests given on the same day, the correlation between the odd and even items of the same tests, while twelve are the intercorrelations between halves of tests given on different days. Since the three tests are regarded as parallel forms each correlation may be regarded as a reliability coefficient of a half test. None of the coefficients has been 'boosted' by the Spearman-Brown formula. Evidence will be advanced later in this paper to show that the three forms used exhibited a high degree of equivalence.

Examination of the matrix of intercorrelations (Table I) between the split-halves of the three parallel forms of the same test shows immediately without the application of Holzinger's *B*-coefficient technique that the correlations between the odd and even items of the same tests are higher than the correlation between the odd or even items of different tests, or between the halves of the tests given on different days.

Table I

	1	2	3	4	5	6
1	—	0.9457	0.9086	0.8937	0.8643	0.8932
2	0.9457	—	0.9081	0.8953	0.8609	0.8969
3	0.9086	0.9081	—	0.9393	0.8806	0.8993
4	0.8937	0.8953	0.9393	—	0.8527	0.8764
5	0.8643	0.8609	0.8806	0.8527	—	0.9278
6	0.8932	0.8969	0.8993	0.8764	0.9278	—

The correlations between halves of the same test have been marked off in Table I by diagonal blocks, and they form non-vanishing tetrads with the other coefficients in the matrix. The correlations between halves of the tests given on different days form tetrad equations whose values do not differ significantly from zero. It is evident, therefore, that it is possible to analyse the present matrix of correlations into a minimum number of one general factor, and three group factors. Since the coefficients in Table I represent the correlations between parallel forms of the same test no specific factor is to be expected. If the tests used had not approximated to a high degree of equivalence, specific factors would have required consideration. The close correspondence of the intercorrelations of the halves of the tests is suggestive that adequate parallelism was secured.

In the present analysis the first factor loadings were estimated by formula (2), and are recorded in the first column of the factor pattern, Table III. The residuals $r_{ij} = r_{ij} - a_i a_j$ were then calculated. The table of residuals after removal of the general factor is given in Table II.

Table II. *First residual correlations*

	1	2	3	4	5	6
1	—	0.0426	-0.0026	0.0024	-0.0010	0.0012
2	0.0426	—	0.0050	0.0036	-0.0047	0.0045
3	-0.0026	0.0050	—	0.0396	0.0071	-0.0012
4	0.0024	0.0036	0.0396	—	-0.0016	0.0044
5	-0.0010	-0.0047	0.0071	-0.0016	—	0.0727
6	0.0012	0.0045	-0.0012	0.0044	0.0727	—

Examination of the first residual matrix indicates that the general factor loadings have described with a high degree of accuracy the majority of the intercorrelations. The residuals r_{12} , r_{34} , r_{56} are, however, considerably larger than the remaining residuals, and indicate the expected tendency of further overlap between variables 1 and 2, 3 and 4, 5 and 6.

The next step in the calculations was to find the error variance of each variable by the formula $e_i^2 = 1 - r_{ii}$, where e_i^2 is the error variance of variable i and r_{ii} is the reliability coefficient of variable i . The loadings of the error factors were thus found, and these are recorded in the staggered e_i column of Table III. In estimating these loadings the odd-even item correlation of each test was taken as r_{ii} , and the assumption made that the odd items of each test had an error variance equal to that of the even items. This is, indeed, a justifiable assumption, and the only one that can be made in the present analysis.

The remaining group factors were then readily calculated by the following simple formula,

$$r_{ib}^2 = 1 - r_{ia}^2 - e_i^2,$$

where r_{ib}^2 is the variance of factor b in test i , r_{ia}^2 the variance of the general factor, and e_i^2 the error factor variance of test i .

Table III. *Factor pattern*

Variable	Factor I <i>a</i>	Factor II <i>b</i>	Factor III <i>c</i>	Factor IV <i>d</i>	Error factor loadings e_i	h^2
1	0.9501	0.2074	—	—	0.2330	0.9457
2	0.9505	0.2054	—	—	0.2330	0.9457
3	0.9591	—	0.1393	—	0.2464	0.9393
4	0.9381	—	0.2435	—	0.2464	0.9393
5	0.9107	—	—	0.3137	0.2687	0.9278
6	0.9389	—	—	0.2152	0.2687	0.9278

The factor pattern of Table III describes with considerable accuracy the original correlation matrix. Some estimation of how closely the final factor pattern accounts for the original correlations is given by examination of the final residuals in Table IV.

Table IV. *Final residual correlations*

	1	2	3	4	5	6
1	—	0.0000	-0.0026	0.0024	-0.0010	0.0012
2	0.0000	—	0.0050	0.0036	-0.0047	0.0045
3	-0.0026	0.0050	—	0.0057	0.0071	-0.0012
4	0.0024	0.0036	0.0057	—	-0.0016	0.0044
5	-0.0010	-0.0047	0.0071	-0.0016	—	0.0052
6	0.0012	0.0045	-0.0012	0.0044	0.0052	—

III. INTERPRETATION OF FACTORS

The factors isolated by the above analysis require interpretation. Close correspondence of the general factor loadings, and also of the group factor loadings, is a good criterion of test form equivalence. Variable 5 (the odd items of M.H.T. 21) manifests the highest degree of inequivalence, but this inequivalence is not sufficiently prominent to introduce a specific factor loading approximating anywhere near significance. The close correspondence of factor loadings as calculated above is a better index of test equivalence than the correspondence of the intercorrelations between split-halves. If the halves of the various tests used are equivalent then the intercorrelations of the halves of the tests given on different days should be equal within the limits of sampling error. The converse, however, does not hold. The fact that the correlations between the halves of tests given on different days are equal is no indication of test equivalence. If A were a test of intelligence and B a test of ability to do arithmetic, and a_1, a_2 are the odd and even items respectively of test A , while b_1, b_2 are the odd and even items respectively of test B , then $r_{a_1b_1}, r_{a_1b_2}, r_{a_2b_1}$, and $r_{a_2b_2}$ could readily be equal, and yet it is obvious that A is a test of different structure from B . What is indicated, however, is that the odd and even items on test A are equivalent, and the odd and even items on test B are equivalent, but the halves of A are not equivalent to the halves of B . Close correspondence of the factor loadings of two forms of a test, when used in the same battery of tests, both parallel forms being applied to the same group of children, is a reliable index of the equivalence of the two forms. In the above analysis the absence of anything approaching a significant specific is demonstrative that good equivalence has been obtained.

The group factors isolated by the above analysis may be termed factors of temporal contiguity, a term first used by Thouless(3). If we could conclude that the function measured were a non-fluctuating one, then these group factors could be interpreted as largely the result of error correlation. If we could conclude that in correlating the split-halves of the same test the errors were uncorrelated, then the group factors could be described as manifestations of the absence of functional variability between those tests having group factors in common. Since, however, it is not unlikely that both the correlation of errors and functional variability are exerting a positive influence on the size of the group factors, and since no method of determining the relative importance of these two influences is at the moment apparent, it is only possible to describe these

factors as factors of temporal contiguity, and to regard them merely as the resultant of those influences that tend to reduce the correspondence between test scores on parallel forms of the same test with increase in the time interval between their applications.

IV. A COMPARISON OF MULTIPLE ORTHOGONAL FACTORS WITH BI-FACTORS

To obtain a comparison between the factors obtained by the above bi-factor analysis, and those obtained by multiple factor analysis(4) the intercorrelations given in Table I between the split-halves of three parallel forms of the same test were analysed by Thurstone's method(5). The largest correlation in each row was used as the diagonal element, and was maintained unchanged throughout the analysis in that it represented a very close approximation to the true communality. It was found that this matrix of correlations could be adequately described in terms of three multiple orthogonal factors instead of four bi-factors. This is in complete correspondence with the findings of Holzinger that four bi-factors can be described in terms of three multiple orthogonal factors. The centroid solution of the Thurstone analysis is given in Table IA.

Table IA				
Loadings of the factors centroid solution				Communality h_i^2
Tests	I	II	III	
1	0.9560	-0.1044	0.1401	0.9445
2	0.9563	-0.1049	0.1498	0.9480
3	0.9602	-0.0690	-0.1217	0.9416
4	0.9465	-0.1297	-0.1687	0.9411
5	0.9320	0.2472	-0.0166	0.9300
6	0.9508	0.1604	0.0111	0.9299

The communalities of the centroid solution are in close agreement with the communalities of the bi-factor solution, and both patterns describe the correlations of the original matrix with a close degree of accuracy.

The factor pattern of Table IA was now rotated to remove negative loadings, and to obtain as many zero loadings as possible, while still maintaining a factor space of three dimensions. This was done by rotating two factors at a time graphically, factors I and II being rotated first, and then factors I and III. Each pair of columns of loadings was post-multiplied by a 2×2 orthogonal matrix representing a rotation of

rectangular axes in two dimensions through a given angle θ . The elements of this orthogonal matrix

$$\begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

were found by regarding the loadings of the test through which the axes were rotated as co-ordinates of a point in a plane, and by these co-ordinates calculating the sine and cosine of the angle of rotation. The rotated factor loadings are given in Table IIA.

Table IIA

Tests	Rotated factor loadings			Communality h_i^2
	I	II	III	
1	0.9229	<i>0.0004</i>	0.3052	0.9449
2	0.9216	<i>-0.0001</i>	0.3149	0.9485
3	0.9687	<i>0.0360</i>	<i>0.0475</i>	0.9419
4	0.9700	<i>-0.0258</i>	<i>0.0000</i>	0.9416
5	0.8881	0.3474	0.1451	0.9304
6	0.9119	0.2631	0.1723	0.9305

The factor pattern of Table IIA is one of a large number that could be obtained by using different angles of rotation. Four of the loadings of factor II, and two of the loadings of factor III are regarded as zero. These loadings are italicized in Table IIA. All other loadings are positive. No system of rotation can produce more than six zeros in this pattern in this three-dimensional factor space. The bi-factor solution describes the observed correlations in terms of four factors and twelve factor loadings. The rotated multiple factor pattern describes the same correlations in terms of three factors and twelve factor loadings.

By the method described by Holzinger in *Student Manual of Factor Analysis* (1) the relationship between the two factor patterns can be found, the relationship being expressed in terms of a set of three linear equations. This involves the reduction of the original tests to as many new variables as there are group factors in the bi-factor solution. In this case the six original tests are expressed in terms of three composite tests z_a , z_b , z_c . The first factor loading of the composite test z_a for both bi-factor and multiple orthogonal patterns is found by adding the first factor loadings of tests 1 and 2, and dividing this sum by the combined standard deviation of these tests. The formula for the combined standard deviation as given by Holzinger for n variables is

$$\sigma_{1+2+3 \dots n} = \sqrt{\{n + 2(r_{12} + r_{13} + r_{14} \dots r_{(n-1),n})\}}.$$

The values in the present case are $\sigma_{1+2} = 1.9729$, $\sigma_{3+4} = 1.9694$, and $\sigma_{5+6} = 1.9635$.

The reduced factor pattern calculated from the bi-factor solution is found to be as follows:

$$\begin{array}{lll} z_a = 0.9635a + 0.2093b & & \overset{h_i^2}{0.9721} \\ z_b = 0.9633a & + 0.1944c & 0.9657 \\ z_c = 0.9420a & + 0.2694d & 0.9599 \end{array}$$

Taking the same composite tests for the multiple factor solution, we obtain the following set of equations:

$$\begin{array}{lll} z_a = 0.9350Z_1 + 0.0001Z_2 + 0.3143Z_3 & \overset{h_i^2}{0.9730} \\ z_b = 0.9844Z_1 + 0.0052Z_2 + 0.0241Z_3 & 0.9697 \\ z_c = 0.9167Z_1 + 0.3109Z_2 + 0.1617Z_3 & 0.9631 \end{array}$$

The communalities in both sets of equations are in close correspondence. The intercorrelations of the reduced tests are given in Table IIIA.

Table IIIA

Multiple			Bi-factor		
	z_a	z_b		z_a	z_b
z_a	—	—	z_a	—	—
z_b	0.9280	—	z_b	0.9281	—
z_c	0.9080	0.9079	z_c	0.9076	0.9074

The two sets of correlations given in Table IIIA are in close agreement and indicate that both patterns are equally good fits of the observed correlations.

By equating these two sets of equations, and solving for Z_1 , Z_2 , and Z_3 , we can attain a set of equations which show the relationship between the two sets of factors by describing the multiple factors in terms of bi-factors. These three equations are found to be

$$\begin{aligned} Z_1 &= 0.9741a + 0.1059b + 0.2144c - 0.0043d, \\ Z_2 &= 0.0712a - 0.3291b - 0.3057c + 0.8880d, \\ Z_3 &= 0.1654a + 0.7226b - 0.6285c - 0.0153d. \end{aligned}$$

The standard deviations of Z_1 , Z_2 , and Z_3 in the above equations approximate to unity, and the intercorrelations of the Z 's approximate to zero.

The relative importance to be attached to each bi-factor in describing

the Z 's may be found by squaring all the values in the above equations obtaining the following:

$$\sigma_{z_1}^2 = 0.9489 \sigma_a^2 + 0.0112 \sigma_b^2 + 0.0460 \sigma_c^2 + 0.0000 \sigma_d^2,$$

$$\sigma_{z_2}^2 = 0.0051 \sigma_a^2 + 0.1083 \sigma_b^2 + 0.0935 \sigma_c^2 + 0.7885 \sigma_d^2,$$

$$\sigma_{z_3}^2 = 0.0274 \sigma_a^2 + 0.5222 \sigma_b^2 + 0.3950 \sigma_c^2 + 0.0002 \sigma_d^2.$$

From these equations it is apparent that nearly all the variance of Z_1 is attributable to the bi-factor a . Z_2 is made up largely of the bi-factor d , while Z_3 is composed largely of the bi-factors b and c .

V. SOME OBSERVATIONS REGARDING THE COMPARISON IN SECTION IV

The above enquiry commenced with the initial hypothesis that factors of temporal contiguity existed, tending to make the intercorrelations between tests given on the same day greater than the intercorrelation between the same tests given on different days. The necessary intercorrelations were calculated, and the postulated factors of temporal contiguity isolated and measured by a bi-factor analysis. It was found that the bi-factor solution furnished a factorial configuration in complete agreement with the postulated psychological hypothesis. The compatibility between the factorial configuration and the psychological hypothesis was sufficient to regard the initial hypothesis as proved.

When we now come to analyse our table of intercorrelations by multiple factor methods we find that an equally accurate mathematical description can be obtained in terms of a pattern of three factors, but no matter what method of rotation is adopted these three factors can never be transformed into a psychologically meaningful configuration within a factor space of three dimensions, a factor space of four dimensions being required before our factor pattern can become compatible with our initial hypothesis. It is of course clear that an orthogonal transformation can in theory be attained capable of rotating the three multiple factors into a psychologically meaningful four factor space. This would involve post-multiplying the factorial matrix of order 6×3 with known elements by an orthogonal matrix of order 3×4 of unknown elements. The estimation of the elements of the orthogonal matrix capable of bringing the multiple factor pattern into agreement with the bi-factor pattern is a matter of considerable mathematical difficulty, and of great arithmetical labour.

In the present example the simplicity of our factor pattern renders the inadequacy of a three dimensional factor space, and the necessity of

an additional space, readily observable. Furthermore the difficulty of attaining a meaningful interpretation of our three rotated multiple factors is also apparent. With more complicated factor patterns, however, this difficulty is not readily observed, and the psychologist has no clue to direct him to the conclusion that his factor pattern must be rotated into additional dimensions to obtain meaningful factors. The assumption is usually made that a minimum number of factors with as many zero loadings as possible is likely to be the most meaningful configuration attainable. In our present example such a configuration has little if any meaning, and it does not seem likely that in more complicated patterns the reduction of the number of factors to a minimum would necessarily lead to the most meaningful solution. Our conclusion is, therefore, that under certain circumstances by reducing the number of factors to a minimum we shall arrive at an invalid interpretation of the mental factors involved in the performance of certain tests, and that under these circumstances bi-factor solutions will tend to more meaningful results than orthogonal solutions.

The fundamental difference between the Thurstone method of obtaining factors and the bi-factor method seems to be this. The former attempts to fit a psychological interpretation to a mathematical hypothesis. The latter attempts to fit a mathematical interpretation to a psychological hypothesis. Since we are primarily interested in proving or disproving psychological hypotheses the bi-factor method would seem, from the point of view of psychology, to be the more valid scientific method, and more likely to produce useful results.

VI. SUMMARY

1. The intercorrelations between the split-halves of three equivalent group tests of intelligence given on different days are analysed by Holzinger's bi-factor method, and factors of 'temporal contiguity' isolated and measured.

2. The existence of factors of 'temporal contiguity' may be due to the absence of the influence of functional variability on the correlations between tests given on the same day, or to the correlation of errors, or both.

3. The existence of factors of 'temporal contiguity' explain why reliability coefficients calculated by the split-half method, and 'boosted' by the Spearman-Brown formula, are usually higher than reliability coefficients obtained by correlating parallel forms.

4. A comparison is made between bi-factor and multiple factor techniques.

5. Reasons and calculations are advanced to show that the reduction of the number of factors to a minimum may under certain circumstances lead to meaningless factors, quite incompatible with a previously established psychological hypothesis.

6. The argument is presented that from the point of view of psychology the fitting of a mathematical interpretation to a psychological hypothesis, rather than the converse, is the more valid scientific method, and likely to lead to more meaningful results.

REFERENCES

- (1) HOLZINGER, KARL J. (1937). *Student Manual of Factor Analysis*. Chicago.
- (2) HOLZINGER, KARL J. & SWINEFORD, FRANCES (1937). "The bi-factor method." *Psychometrika*, II, 41-54.
- (3) THOULESS, ROBERT H. (1936). "Test unreliability and function fluctuation." *Brit. J. Psychol.* XXVI, 325-43.
- (4) HOLZINGER, KARL J. (1938). "Relationship between three multiple orthogonal factors and four bi-factors." *J. Educ. Psychol.* XXIX, 513-19.
- (5) THURSTONE, L. L. (1935). *The Vectors of Mind*. Chicago.

(Manuscript received 21 January 1940)

PUBLICATIONS RECENTLY RECEIVED

The Measurement of Abilities. By P. E. VERNON. University of London Press. 1940. Pp. 308. 10s. 6d. net.

This is an admirable book which should relieve teachers of educational psychology of the difficulty of finding a suitable text-book for their classes. It provides an interesting and clearly written account of the use of statistical methods, of mental testing, and of examining. The account of statistical methods should be valuable to those students who find Fisher too hard, but who want an account of the matter which is practical and sound, and who also prefer one that is illustrated from the field of educational psychology.

Criticisms of this part of the book must be mainly on questions of relative emphasis on which there would be considerable disagreement between different educational psychologists. I should have preferred to see a fuller account of the method of testing significance in small samples. It is true that, in educational psychology, one is generally dealing with relatively large samples, but not always, and ability to deal adequately with small samples is a useful part of the equipment of every psychological statistician. In experimental psychology, it is, of course, essential.

Much space is rightly devoted to the correlation coefficient since this is the method which has been most widely used in educational psychology. Dr Vernon also very properly points out that adherence to the correlation coefficient commonly leads to neglect of more appropriate methods of dealing with particular problems, such methods, for example, as the analysis of variance. It might have been better to have given a fuller description of the analysis of variance, not only because it is a method which should very often be used, but because it is the expression of a way of thinking which is essential to the proper understanding of the use of correlation coefficients. Much of the blind error in the use of correlation coefficients in the past (such as Winfield's determination of the correlation between the intelligence coefficients of pairs of twins from a single orphanage) would have been avoided if investigators had realized that, in calculating a correlation coefficient, one is comparing variances.

There is an interesting and useful chapter on factor analysis. In dealing with the difficult problem of multiple factor analysis, the author has adopted the wise course of explaining what multiple factor analysis is getting at without giving instructions in the actual methods of carrying it out. I do not think that there is anywhere else where this is done for the student. If he wishes to learn the actual techniques, he will be in a much better position for doing so after he has understood Dr Vernon's chapter.

Dr Vernon argues again the case against the traditional examination as an adequate measuring instrument. It might have been better to omit much of the anecdotal evidence, such as the story of the examiner who wrote for his own guidance a model set of answers which were marked variably and sometimes severely by his colleagues. It is difficult to know how far such selected evidence is truly representative. Even the more respectable evidence of correlation coefficients may often be symptomatic of bad examining rather than of a necessary defect in the examination. Independent marking of university examinations by competent examiners is found to give considerably higher correlations than those ordinarily published. On the other hand, it may also be pointed out that the matter may be worse than correlation coefficients indicate since these indicate only the reliability of independent examiners' marks, not their validity. Unless papers are independently marked from typed copies, we cannot know how much of such agreement as there is may be due to both examiners being influenced by handwriting and neatness rather than by the candidates' knowledge of the subject. The main practical problem for the examiner is to use such techniques as will reduce the unreliability of his marking as much as possible. On this point, Dr Vernon gives practical guidance which would much improve marking if it were followed.

Towards the end of the book, there is a new-type examination on its subject-matter which the reader will find it amusing to carry out himself. There is a key to right answers at the end. Perhaps in a future edition the author will add norms.

This is a thoroughly useful book which fills a gap of which teachers of educational psychology have been long conscious. R. H. T.

An Annotated Bibliography of Mental Tests and Scales. By CHARLES K. A. WANG. Peiping, China: Catholic University Press. 1939. Pp. vi + 725.

More than 1770 references are given and their contents briefly described. The general topics more or less covered in this first volume are: the Measurement of Mental Capacity; the Measurement of Personality and Character; and the Measurement of Vocational Aptitudes and Abilities. A list is also included of the publishers and manufacturers of test material.

Whether the preparation of this kind of bibliography is genuinely worth the labour that must be lavished upon it is perhaps doubtful, but it is certain that Professor Wang and his collaborators have accomplished the job itself in an exceedingly satisfactory manner. The annotations are brief but clear, and where they cannot be first-hand this is indicated. Anybody turning over the pages of this book will probably get a very strong impression that research workers in this field have been more anxious to invent tests than to do anything else, and also that the number of tests is altogether out of proportion to their originality.

It is probable that the bibliography will be most useful to the student who wishes to get a quick specialized survey and then to follow this up by study of original papers in a more limited field.

General Psychology. By L. E. COLE. London: McGraw Hill Publishing Co., Ltd. 1939. Pp. xii + 688. 21s. net.

This is a comprehensive, judicious, and lucid general text-book. Prof. Cole reveals a wide knowledge of his subject-matter, a healthy bias towards treating the subjective in an objective a fashion as possible, and a great ability to keep his head amidst the wildest swirl of theories. On the whole, the arrangement of this text-book follows a fairly well established modern pattern. There are the initial physiological data, very well presented, the early stress on general and particular biological development. Then come the psychological parts proper: emotion, motivation, learning, and, as compared with the traditional type of text-book, perception at a late stage. There are both fuller and better discussions of thinking and reasoning than have been customary, and the final chapters deal with intelligence and personality. The book is undeniably a good one for the intermediate and advanced student who can read with some independence of mind. He may find parts of it over long, but he will get a good picture of current psychology and may well be stimulated towards an attitude neither over-critical nor merely submissive.

Representative Psychologists. By A. C. WESTERHOF. Maryland: Pilot Publishing Company. 1938. Pp. vi + 119.

The title is not a very good one, for the book is about psychological theories rather than about psychologists. The theories selected are: Personalistic, mainly represented by William Stern; Gestalt, represented by Köhler and Koffka; Mental Measurement, represented by Terman; Purposive Behaviorism, represented by Tolman; Psycho-analytic, represented by Freud, Jung and Adler; and Hormic or Teleologic, represented by August Pauly and William McDougall. The author, while very generally critical, is more sympathetic to the hormic than to any other point of view. The essay

is worth reading, but on the whole tends to exaggerate the differences between one group and another in modern psychology. The fundamental problem to the author remains the apparent opposition between mechanistic and teleological views, and he calls for a reorientation in regard to this based upon a re-examination of the "nature of thinking itself". However, there is not very much indication of how the re-examination is to be carried out.

The Individual and His Society. By ABRAM KARDINER. New York: Columbia University Press (London: Humphrey Milford). 1939. Pp. xxvi+503. 22s. net.

This book is a co-operative effort. It contains a fair amount of original sociological and anthropological observation by Dr Ralph Linton, dealing with Marquesan culture and the Tanala of Madagascar, two peoples who look to be conveniently chosen for their wide differences of social organization and character. It also contains more interpretation by Dr Abram Kardiner, who disclaims any special knowledge of anthropology and whose approach to psychology is very strongly by way of a study of the pathological. This has certain disadvantages: the descriptive parts of the volume have to be highly condensed and selected and the interpretative parts often give the impression of being at several removes from the facts, as if the author had achieved a high degree of theoretic formulation before he attempted his concrete task. Nevertheless, the book is a highly interesting one and the authors themselves very frankly state its shortcomings and incompleteness. It falls in general into line with many other recent books in strongly using the notion of group personality. It insists, both that the psychologist should take account of this, and also that the anthropologist must learn to reckon more shrewdly with the functions of the individual in society. And it attempts to work out these and related ideas, not in general terms merely, but with special reference to the two cultures which are studied in some detail. Incidentally, a number of very interesting observations are made about the relations of culture and neurosis, and there are a few valuable and original observations about special cases of the contact and spread of cultural features. It is a pity that the English price of the book is so high.

Time-Budgets of Human Behaviour. By P. A. SOROKIN and C. Q. BERGER. Harvard University Press (London: Humphrey Milford). 1939. Pp. vi+204. 10s. 6d. net.

On the basis that man is an active animal the authors set out to attempt to establish what proportion of time is normally allotted to a large number of everyday activities in certain communities. A considerable amount of the work was done in Soviet Russia and the rest in the U.S.A. The result is an exceedingly interesting book. Various methods were exploited, questionnaires, diaries, immediate records of one kind and another and most of the common things that men can do have been more or less fully covered. The book has four essential parts, dealing with (a) the structure and composition of human behaviour in terms of the relative expenditure of time on its leading component parts; (b) the motives that lie behind the main overt activities; (c) a comparative study of social and non-social activities; and (d) comments on the predictability of human activities and social processes. The authors remain somewhat unimpressed by the modern claim to foresee and plan ahead the activities of man, for they find that most of the primary activities (in terms of time) are relatively inelastic. The evidence is presented most clearly and in a highly condensed manner and the comments that are made are straightforward and definite. It is to be hoped that the book will set other people working on similar problems in other communities. Time may be no great criterion of value, but undeniably it has a lot to do with what is possible in the way of human development.

Psychological Foundations of Personality. By LOUIS P. THORPE. London: McGraw Hill Publishing Co., Ltd. 1938. Pp. xvi + 602. 21s. net.

Prof. Thorpe describes his book as a Guide for Students and Teachers. The description is accurate, but at the same time the book has an interest and a value that run beyond the preparation for examinations. It is clearly planned. The author discusses first the current general views about the basis of personality; then considers the biological foundations of personality; proceeds to a fairly wide survey of educability, largely in terms of the training of intelligence; next considers dynamic and emotional factors, and the significance of special "traits"; examines both disturbances of personality and its integration; appraises certain physiological theories; and finally reaches some tentative conclusions with regard to problems of measurement in this field. The actual writing, though in no way distinguished, is usually clear, though in places definitely long-winded. The work surveyed is very large in quantity and the surveys are judicious and within their limits very fair to the authors dealt with. While the volume is not exciting, it is informative to a high degree.

Secrets of the Mind. By C. E. WAGER. London: Watts and Co. 1940. Pp. 159. 5s. net.

The 'secrets' are, of course, mainly concerned with those mechanisms which are now usually called Freudian. Whether 'secret' is the correct word to use about them now, when more has been written concerning them than concerning most other things in psychology, is perhaps slightly dubious. Anyhow, the secrets are in this book partially and as it were discreetly unveiled. Most of the chapters are short and read somewhat like lecture notes. They deal with questions connected with family life, with marriage, with motives, with religion, with neuroses and finally, and very briefly, with dictators, who appear to be more or less explained by being called wrong super-egos, or substitute super-egos (it is not quite clear which). Taking it all through, the book may well be useful if it is to be read in conjunction with a well-run critical course; but as reading for the person who is not much trained and has little access to anything else it is not to be recommended. Its sincerity shines all through, but it is neither as clear in its exposition nor as critical in its outlook as could be desired.

Hopousia, or, the Sexual and Economic Foundations of a New Society. By J. D. UNWIN. London: Geo. Allen and Unwin. 1940. Pp. 475. 21s. net.

Any review of this work would be difficult, but an adequate short notice is impossible. Its scope and aim may be indicated, and the reader who feels attracted be left to form his own opinions. Hopousia, we learn on p. 119, is a "hypothetical or experimental society" which is so built that it can display "the greatest possible energy". Precisely how this criterion can be applied is not very clear, but it does seem clear that in many details Hopousia would contain compromises. For example, *alpha* and *beta* marriages are contemplated, the former fixed, binding, unbreakable, the latter shifting, occasionally transitory; but it is the former alone that seem consistent with the author's notion of greatest possible energy. On the whole, the bulk of the work is concerned with economic considerations, and the bulk of these with exchange problems. Beyond a doubt an extremely large number of acute critical comments are made upon institutions and customs of present-day society, but the constructive side of the book, though it is much to the fore, is not as fully worked out. Actually, Unwin did not finish very much of the publication, and a large part of it should not be regarded as a final presentation of his views. Aldous Huxley writes a long and commendatory preface.

A Study of a Group of Children of Exceptionally High Intelligence Quotient in Situations Partaking of the Nature of Suggestion. By RACHEL M. SIMMONS. New York: Teachers' College, Columbia University, Contributions to Education, No. 788. 1940. Pp. x+112. \$1.60.

If a reader thinks that the extraordinarily cumbrous title of this monograph indicates a mind careful and anxious to be extremely accurate he will be right. If he considers also that it may suggest a mind cautious to the extent of being uninteresting he will be wrong. The monograph is in fact both illuminating and attractive. To begin with, it gives a very good survey of earlier work on suggestibility in young people, and then it goes on to describe clearly and to discuss a number of experimental investigations carried out with two carefully paired groups of intelligent children and less intelligent ones. The procedures used were the Otis tests I and II and other Street picture tests, progressive weights and the Chevreul pendulum. Throughout, the more intelligent children showed smaller tendency to accept suggestions, and greater tendency to resist suggestions in a positive way. The author interprets these results most reasonably and with extreme care. She is very desirous of not running beyond her own particular groups and her own arranged situations; nevertheless, she does seem to have provided convincing experimental evidence on the basis of which it would be safe to venture wider generalizations.

Marriage and the Family. By RAY E. BABER. London: McGraw Hill Publishing Co., Ltd. 1939. Pp. xii+656. 24s. net.

Here is a general text-book on very nearly all aspects of family life or lack of life. And if a text-book on these subjects is desired this is an excellent one. It covers a vast range, historical, economic, physiological (but not very much of this), psychological, and social, and covers this range well. It is full of information, and some of this, especially that concerning earlier and current American custom and institution, is particularly interesting to readers in this country. On the whole, the family as a social institution of fundamental value comes well out of the survey, and the author, though he meets his problems with a critically open mind, is definitely on the side of what he calls the "conservation of family values". Except for the specialist the book is on the long side, but no psychological student can fail to be interested and enlightened by many parts of it, and general psychological libraries would do well to place a copy on their shelves.

The Rape of the Masses. By SERGE CHAKOTIN. London: George Routledge and Sons. 1940. Pp. xviii+299. 7s. 6d. net.

The Rape of the Masses has enjoyed great popularity in France and is now translated into English from that language, and translated very well. It is topical psychology with a vengeance. Its sole subject is propaganda and its sole explanation conditioned reflexes. Only instincts and especially "no. 1 instinct, the combative instinct" take the place of reflexes. Whatever may be thought about the theory, and it is likely that most readers with psychological training will find it continually exaggerated and over-strained, there is no doubt that much of the author's factual evidence is of absorbing interest.

Guiding Human Misfits. By ALEXANDRA ADLER. London: Faber and Faber. 1939. Pp. 128. 5s. net.

While there is nothing very original or striking about this book it is written in a clear and straightforward manner. Dr Alexandra Adler follows her father, to whom

she has dedicated this book, very closely, and her present work consists largely in briefly described case-illustrations of various difficulties which have been successfully treated along the lines of Individual Psychology.

Alfred Adler: the Man and His Work. By HERTHA ORGLER. London: C. W. Daniel Company, Ltd. 1939. Pp. 241. 8s. 6d. net.

Alfred Adler. By PHYLLIS BOTTOME. London: Faber and Faber. 1939. Pp. 386. 10s. 6d. net.

Both books have sub-titles: that of the first is *Triumph over the Inferiority Complex*, and that of the second is *Apostle of Freedom*. Both books are what may be called 'praise biographies'. Mrs Bottome's book is perhaps the better written and certainly gives more personal details of its hero in a more consecutive and orderly fashion. Mrs Orgler is perhaps a little more concerned to extol the soundness of Adler's principles and the efficacy of his methods. This she does partly by descriptions of her own applications of his ideas in medical practice. The many admirers of Adler's work will find much to attract them in both of these biographies.

How You Grow Up. By VICTOR CHAMBERLAIN. London: Gerald Howe, Ltd. 1939. Pp. v + 121. 2s. 6d. net.

It is difficult to see what possible useful purpose a book like this can serve. It is called "an Introduction to Psychology for young people and their parents". It is said to be "suitable for all ages from thirteen". In fact, it consists of a number of exceedingly brief, exceedingly general and exceedingly inadequate chapters on very widely selected topics. Any intelligent thirteen-year-old, if he is in need of anything specifically in the name of psychology at all, deserves something better than this. His parents possibly may be more interested, but they will not be able to learn anything much.

Children in the Cinema. By RICHARD FORD. London: Geo. Allen and Unwin. 1939. Pp. viii + 232.

It is estimated that in normal times very nearly five million children attend cinema performances each week in this country. Here is a survey of and a tentative answer to a great many problems of social and psychological interest which result. The author apologizes for his lack of expert knowledge of child psychology. No such apology was needed. The book is most admirably clear, very much to the point, full of interest, and it will suggest to the psychologist who cares for the world outside as well as that inside the laboratory a number of very worth-while lines of investigation. In fact, this is a book to be studied and also to be enjoyed.

Sleep and Wakefulness. By NATHANIEL KLEITMAN. University of Chicago Press (Cambridge University Press). 1939. Pp. xi + 638. 30s. net.

This is a very excellent book. It passes in review, swiftly but lucidly, an immense amount of observational and experimental research concerning the physiological, pharmacological and psychological characteristics and functions of the sleep-wakefulness periodicity. It includes a most interesting chapter on the experimental deprivation of sleep, which shows how, as lack of sleep increases, the pain threshold is usually lowered, irritability grows, hallucinations may appear, and in general the signs indi-

cate growing fatigue of the higher levels in the central nervous system. Sleep-like states, including hypnosis, are judiciously considered, and there is much about abnormal sleep and wakefulness. Very likely the general reader may find the final part of the volume, which deals with theories of sleep, the most interesting. The author critically considers neural and humoral theories, the sleep-centre problem, and an evolutionary theory which treats sleep and wakefulness as natural alternating phases in a biological cycle. Dr Kleitman's adverse criticism of the widely accepted inhibition theory of Pavlov is incisive and seems entirely convincing. The author himself favours a form of the evolutionary theory, but freely admits that there are many gaps still to be filled.

The Psychology of Art. By ROBERT MORRIS OGDEN. New York: Charles Scribner's Sons. 1938. Pp. xviii + 291. \$2.50.

Most of the traditional problems of psychological aesthetics are discussed in this book. After a general introduction dealing with the wide field of all the Arts, Prof. Ogden discusses in turn Music, Poetry, Visual Art, combinations of the different modes of artistic expression and, rather briefly, aesthetic education. He looks everywhere, not for items of expression and brief incidents of enjoyment, but for patterns which may be exceedingly complex in themselves and for 'enjoyed' responses which may contain a great many components welded together so that their secret lies in their unity, not in their units. "Musical tones determine among themselves patterns of melody and harmony which, when apprehended, open a wide prospect of aesthetic enjoyment." "Great literary art reduces itself to a verbal pattern which we can fully enjoy." Visual and the combined arts display the same characters. Yet Prof. Ogden does not despise the study of the component detail, but on the contrary considers it essential and devotes much attention to it. The whole treatment cannot be regarded as particularly original, but it is vigorous, clear and pleasant to read. If there are key-words, they are 'orderliness', 'seemliness', 'balance'. The volume is exceedingly well produced and illustrated.

The Psychology of Music. By CARL E. SEASHORE. London: McGraw Hill Publishing Co., Ltd. 1938. Pp. xix + 408. 24s. net.

There is no need to attempt to praise a book on this subject by Prof. Seashore. His work is already so widely known, so warmly appreciated in all the civilized countries that it is enough to record publication. The volume consists of a very considerable number of essays, some of which have been published before but are not easily accessible, on a great number of different topics both of technical and of general interest, connected with music. On all of them Dr Seashore writes with great knowledge, *verve* and distinction, and his book is sure of a wide welcome.

The Ultra-perceptive Faculty. By J. HETTINGER. London: Rider and Co. 1940. 12s. 6d. net.

This is an experimental investigation of what in psychical research is called 'psychometry'—the alleged power of a medium to give information about the owner of an object on which the medium has concentrated. The general principle of the methods used has been to present the owner of the object with the statements made by the medium about him and with a number of control statements not made about him and to determine whether he accepts a significantly greater number of the statements which are meant to apply to him. There is considerable variation in detail of the method, but it seems clear that there is definite evidence of the existence of this power amongst professional mediums when the conditions of the experiment are entirely adequate. The statistical treatment is also adequate although somewhat stereotyped and insensitive.

The fact of the power seems to be proved, but there are many interesting statements about it, of which no numerical proof is given, under the head of 'further observations' at the end of each chapter. Some of the conclusions in the latter part of the book are very questionable since they seem to assume that every time a statement is 'accepted' by the person to whom it refers, genuine ultra-perceptive cognition is at work, whereas the number of statements accepted in the control group suggests that only a small proportion of the correct statements are genuinely cognized and that the remainder may be true by mere chance.

R. H. T.

Diseases of the Nervous System. By W. RUSSELL BRAIN. Second Edition.
Oxford Medical Publications (London: Humphrey Milford). 1940.
Pp. xx + 950. 30s. net.

The second edition of this well-known and valuable text-book has been widely and carefully revised. It contains now a chapter on the Psychological Manifestations of Organic Nervous Disease, discussions of modern electro-encephalography, references to the effects of vitamin deficiencies in relation to nervous disorder, and brief accounts of the progress made in the study of neurotrophic viruses. A few sections, less immediately neurological in character, have been omitted.

The psychological chapter, though it is clear and accurate, is necessarily brief and can claim no originality. Moreover, the tests and methods of investigation referred to are hardly given sufficient detail to guide their application. However, all the necessary references are appended from which readers who desire it may obtain more detailed information. The volume should continue its successful career.